DOCUMENT RESUME

ED 464 344 CS 510 951

TITLE	Evaluation of DC Reads: Year 2 Final Report.
INSTITUTION	Macro International, Inc., Silver Spring, MD.
SPONS AGENCY	Corporation for National Service, Washington, DC.
PUB DATE	2000-01-31
NOTE	87p.; Appendix C, which contains 13 survey instruments, is neither included in the on-line version nor the ERIC version.
AVAILABLE FROM	For full text: http://www.cns.gov/areads/about/evaluation.html.
PUB TYPE	Reports - Evaluative (142)
EDRS PRICE	MF01/PC04 Plus Postage.
DESCRIPTORS	*Outcomes of Education; *Partnerships in Education; Primary Education; *Program Evaluation; Program Implementation; *Reading Achievement; Student Needs; *Tutorial Programs
IDENTIFIERS	Corporation for National Service; District of Columbia

ABSTRACT

The DC Reads program was first implemented during the 1997-1998 school year as a partnership among the Corporation for National Service; District of Columbia Public Schools; Communities in Schools (CIS), a nonprofit corporation; and six local universities, with Georgetown University acting as lead. During the 1998-1999 school year, it expanded to include two more city-based universities and more nonprofit community organizations that coordinate the delivery of testing services. The Year 2 evaluation of DC Reads focused on the program operated in 16 schools by CIS during the 1998-1999 school year. Across its 16 sites (4 of which served as primary study sites) DC Reads provided tutoring to approximately 504 students during the school year--most were in second grade. Students providing tutoring numbered 340; in the 4 study schools, 136 students received tutoring from 74 tutors. DC Reads tutors were mostly Federal Work-Study students attending partner universities, and they used a curriculum and instructional materials purchased by the partners and adapted. Most schools used the Book Partners curriculum while a few schools used the Reading One-to-One curriculum. The Year 2 evaluation concentrated on assessing student outcomes. Findings include: significant gains on 5 of 8 tests in the Reading Performance Battery, with dramatic gains in phonemic awareness; lowest performers progressing at the fastest rate; and nearly twice as much gain for tutored students compared to nontutored students on the Stanford Achievement Test. Effective practices and lessons learned (solutions) are: use of research-based elements to produce reading achievement; well-structured tutoring sessions in which the content and delivery of instruction are carefully planned; close coordination with school personnel; intensive, ongoing training and supervision for tutors; frequent, regular tutoring sessions; careful evaluation, assessment, monitoring, and reinforcement of progress; access to training and technical assistance resources; and engendering positive, caring relationships among student, staff, and tutors. Students who received tutoring made significant gains in reading, and overall the program experienced considerable success in its second year. Appended are: methodology, sources of information, and outcomes. (NKA)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 464 344

EVALUATION OF DC READS

YEAR 2
FINAL REPORT

Submitted to:
Corporation for National Service

Submitted by:
Macro International Inc.

January 31, 2000

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☐ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

CS 510 951

Table of Contents

Page

DC Reads Year 2 Evaluation Executive Summary i

I. Chapter I: Student Outcomes I-1

A. Outcome Study Overview I-1

B. Data Sources for the Outcome Study I-2

C. Changes on the Reading Performance Battery: Comparison of Wave 1 and Wave 3 ..... I-4

D. Multivariate Analysis of the Effect of Tutoring on Reading Improvement I-8

E. Students' Stanford Achievement Test Reading Scores I-11

F. Students' Grades in Reading I-13

G. Summary of the Outcome Study I-13

II. Chapter II: Program Description II-1

A. The DC Reads Partners II-1

B. Program Operations II-1

C. How Children Learn to Read II-2

D. Curriculum and Instructional Materials II-2

E. Session Activities II-4

III. Effective Practices and Lessons Learned III-1

A. Overview of the Implementation Assessment and Investigation of Program Achievements III-1

B. The Effective Practices Model III-2

C. Eight Key Components of Reading Tutoring Programs III-3

IV. Conclusion IV-1

A. Achievements Related to Student Outcomes IV-1

B. Implications for Other Tutoring Programs Focused on Reading Improvement IV-2

C. Next Steps for DC Reads IV-4

BEST COPY AVAILABLE

Appendixes

Appendix A: Methodology **A-1**
A. Evaluation Components A-1
B. Study Schools and Students A-1
C. Assessment Schedule for Students Participating in the Outcome Study A-2
D. Data Sources A-2

Appendix B: Sources of Information **B-1**
A. Sources of Data Supporting Effective Practices and Lessons Learned B-1

Appendix C: Instruments **C-1**
– Teacher Form: Selection of Students C-1
– DC Reads Book Partners Progress Log C-2
– AmeriCorps*VISTA Form: Log of Tutored Students C-3
– AmeriCorps*VISTA Form: List of Tutors C-4
– Book Partners Tutor Observation Protocol C-5
– Reading One-to-One Tutor Observation Protocol C-6
– AmeriCorps*VISTA Survey Questionnaire C-7
– Tutor Survey Questionnaire C-8
– Project Site Director Survey Questionnaire C-9
– Campus Coordinator Survey Questionnaire C-10
– Principal Survey Questionnaire C-11
– Teacher Survey Questionnaire C-12
– Parent Survey Questionnaire C-13

Appendix D: Outcomes[1] **D-1**
A. Changes in Student Reading Performance: Comparison of Wave 1 and Wave 3 ..... D-1
B. Stanford Achievement Test Results D-4
C. Correlation of Reading Performance Battery with SAT 9 Reading Subtests D-4
D. Attitudes Toward Reading D-4
E. Language Arts Grades D-9
F. School Attendance D-9
G. Multivariate Analysis of Student Reading Improvements D-10
H. Summary of Findings D-22

[1] Appendix D is a technical appendix, which includes numerous tables related to the outcome study. It is intended for a technical audience and has not been included in all copies distributed. If Appendix D is not included in your copy and you would like to receive it, please contact Dr. Susan Labin in the Corporation's Office of Evaluation.

List of Figures and Tables

Figures

In Report

Figure 1: Reading Performance Battery Tests I-2
Figure 2: Reading Performance Battery Tests, Wave 1 Wave 3 Comparison I-5
Figure 3: NCE SAT 9 Total Reading Scores I-12

Tables

In Report

Table 1: Reading Performance Battery–Means for Wave 1, Wave 2, and Wave 3 ..... I-4
Table 2: Woodcock Reading Mastery Tests–Mean Grade Equivalent Scores I-5
Table 3: Change in Tutored, Tested Students From Wave 1 to Wave 3 ..... I-6
Table 4: Change on Durrell Analysis of Reading Difficulty I-7
Table 5: Comparison of Grades of Tutored and Nontutored Students I-13
Table 6: Reading Stages, Curriculum Materials, and Outcome Measures II-5

In Appendixes

Appendix B: Sources of Information

Table B-1: Supporting Effective Practices and Lessons Learned by Key Program Components B-1

Tables

IN APPENDIX D: OUTCOMES

Table D-1: Change in Tutored, Tested Students on Reading Performance Battery from Wave 1 to Wave 3 ..... D-2

Table D-2: Durrell Analysis of Reading Difficulty D-3

Table D-3: Stanford Achievement Test (SAT 9) Normal Curve Equivalent Scores D-3

Table D-4: Continuous Data Correlations–Fall '98 SAT 9 and Wave 1 Individual Tests D-7

Table D-5: Ordinal Data Correlations–Fall '98 SAT 9 and Wave 1 Individual Tests D-8

Table D-6: Continuous Data Correlations–Spring '99 SAT 9 and Wave 3 Individual Tests D-9

Table D-7: Ordinal Data Correlations–Spring '99 SAT 9 and Wave 3 Individual Tests D-10

Table D-8: First and Fourth Quarter Language Arts Grades D-15

Table D-9: Comparison of Absenteeism and Tardiness of Tutored and Nontutored Students D-16

Table D-10: Description of Variables Used in HLM Analyses D-18

Table D-11: Reading Performance Battery, Means for Wave 1–Wave 2 and Wave 3 ..... D-19

Table D-12: Woodcock Reading Performance Battery Tests–Means for Wave 1, Wave 2, and Wave 3–Grade Equivalent Scores D-20

Table D-13: Description of the Independent Variables D-21

Table D-14: Linear Model of Growth in Percentage of Capital Letters D-23

Table D-15: Linear Model of Growth in Percentage of Lower Case Letters D-23

Table D-16: Linear Model of Growth in Phonemic Awareness D-24

Table D-17: Quadratic Model of Word identification D-25

Table D-18: Quadratic Model of Growth in Word Identification–Effects of Reading Activities D-26

Table D-19: Linear Model of Growth in Word Attack D-26

Table D-20: Linear Model of Growth in Passage Comprehension D-27

Table D-21: Linear Model of Growth in Passage Comprehension–Effects of Total Days in the DC Reads Program and Percentage of Reading Activities D-28

DC Reads Year 2 Evaluation
Executive Summary

A. Background

The DC Reads program was first implemented during the 1997-1998 school year as a partnership between the Corporation for National Service (Corporation); District of Columbia Public Schools (DCPS); Communities in Schools (CIS), a nonprofit organization; and six local universities (American, Catholic, George Washington, Howard, and Trinity), with Georgetown University acting as lead. During the 1998-1999 school year, the program expanded to include two more city-based universities, Southeastern and the University of the District of Columbia. It also expanded to include more nonprofit community organizations that coordinate the delivery of tutoring services.

This Year 2 evaluation of DC Reads focused on the programs operated in 16 schools by CIS during the 1998-1999 school year. Four of the schools served as primary study sites for the evaluation.

Across its 16 sites, CIS DC Reads provided tutoring to approximately 504 students during the 1998-1999 school year. Most tutored students were in second grade. A total of 340 tutors provided tutoring to those students. In the four study schools, 136 students received tutoring from 74 tutors.

DC Reads tutors, most of whom were Federal Work-Study students attending partner universities, provided tutoring to first- through third-grade students twice a week up to an hour each time. Tutors used a curriculum and instructional materials that the partners purchased and adapted. Students who participated in DC Reads were those identified by school staff as having low reading skills and being most in need of tutoring.

At the beginning of the 1998-1999 school year, most schools used the Book Partners curriculum while a few schools used the Reading One-to-One curriculum. At the start of the second semester, a new blended curriculum incorporating both was adopted in all schools. The blended curriculum enabled tutors to meet the needs of a wide range of students who received tutoring, including those with very low reading skills.

B. Outcomes

The Year 2 evaluation concentrated on assessing student outcomes, although it also included a program implementation assessment and an investigation of program achievements.

1. Significant Gains on 5 of 8 Tests in the Reading Performance Battery, With Dramatic Gains in Phonemic Awareness

The purpose of the outcome assessment was to determine whether tutored students' reading skills improved after participating in the DC Reads tutoring program. The central element of the outcome study was the Reading Performance Battery, which consisted of eight individually-administered reading assessments. Most tests in the battery were administered three times. Between the fall and spring testing (Wave 1 and Wave 3), tutored students demonstrated statistically significant improvement on five of the eight measures in the Reading Performance Battery. Students improved in letter identification, phonemic awareness, word attack, and oral reading fluency and comprehension. Their most dramatic gains were in phonemic awareness, the matching of sounds to their letter symbols. This is considered to be a fundamental skill, prerequisite to the development of more advanced reading skills.

There were no statistically significant declines in scores, even on the subtests that are adjusted for time on a monthly basis. On six of the eight tests, the standard deviation of scores decreased, a change indicating that students performed not only better but also more consistently as a group in Wave 3 testing. This means that fewer students were falling far behind.

2. Lowest Performers Progressing at the Fastest Rate

By using Hierarchical Linear Modeling (HLM), a statistical analysis technique, the evaluators determined there was significant variation among tutored students in their initial scores on measures of phonemic awareness, word identification, and passage comprehension. For each of these outcome measures, students who had lower scores tended to grow at a faster rate than those with higher scores. This indicates that the lowest-performing students were "catching up" in some areas. HLM also indicated that students who practiced reading aloud more frequently during tutoring improved at a higher rate than students who did so less frequently.

3. Nearly Twice as Much Gain for Tutored Students Compared to Nontutored Students on the Stanford Achievement Test

Tutored students made greater gains on the Stanford Achievement Test (SAT 9) reading tests than did nontutored students. On Normal Curve Equivalent scores (which compare students to all other students in the nation at the beginning and end of the school year), tutored students gained nearly twice as much as nontutored students. Also, 69 percent of tutored students improved on Total Reading, while just 60 percent of their nontutored peers demonstrated improvement.

Although tutored students performed significantly below nontutored students in both fall and spring testing, their SAT 9 score improvements offer evidence that tutored students made significant gains above and beyond what would be attributable to the reading instruction received in their classroom.

C. Effective Practices and Lessons Learned

The Corporation defines ***effective practices*** as the activities and methods a program engages in to produce positive results. ***Lessons learned*** are the solutions implemented in the program or envisioned as improvements for future programs, often developed in response to the challenges the program confronted. This study uses the Corporation's previously identified[1] eight principles of high-quality tutoring programs to present the effective practices and lessons learned from the Year 2 DC Reads program.

Macro collected data related to program implementation and achievements by means of interviews, observations of tutoring sessions, various program recordkeeping logs, and survey questionnaires.

1. Use of Research-based Elements to Produce Reading Achievement

The CIS DC Reads program used Book Partners and Reading One-to-One, two well-tested reading curricula developed at major universities to improve the reading abilities of low-performing students. (Book Partners is a DC Reads adaptation of Book Buddies, which was developed at the University of Virginia; Reading One-to-One was developed at the University of Texas by George Farkas.) The selection and use of these curricula comprised the effective practice component of the program while the adaptation of the model (e.g., blended curriculum) to better meet the needs of students reflected a lesson learned. A second lesson learned was that the blended curriculum still did not meet the needs of all children in the program, particularly those who had severe reading difficulties and those who progressed but did not progress enough to be graduated from the program.

2. Well-structured Tutoring Sessions in Which the Content and Delivery of Instruction Are Carefully Planned

The CIS DC Reads tutors used formal written lesson plans to guide each tutoring session. The curriculum materials were geared to a wide range of reading abilities, and there was a corresponding lesson plan for each book. One lesson learned was that the large manual developed to guide the lessons in one curriculum was unwieldy and was not well received by tutors. The second lesson learned was that the increased range of materials and clear lesson plans relieved tutors of having to invent their own activities during tutoring sessions.

---

[1] The principles are described in the Corporation's *Principles and Key Components for High-Quality America Reads and National Service Program Initiatives* and are grounded in the literature on reading research.

3. Close Coordination With Schools, School Administration, and the Classroom and/or Reading Teacher

Related to close coordination, the CIS DC Reads program involved teachers in selecting the students for the tutoring. Second, the main DCPS office was involved in the program and encouraged principals to take an active role in supporting the program. The presence of AmeriCorps*VISTA members as program coordinators at each participating school provided a point person for school staff. Providing information about the program to school staff at the beginning of the school year was also an important effective practice.

Improving access to facilities and equipment, increasing communication with teachers, increasing access to the school-based reading specialists, and recognizing how cultural (organization and ethnicity/race) differences impact communication comprise the lessons learned related to close coordination.

4. Intensive and Ongoing Training and Supervision For Tutors

For this program element DC Reads had two effective practices. One effective practice was hiring an individual with a background in reading to coordinate all DC Reads training activities. Another effective practice was to have designated individuals, in this instance AmeriCorps*VISTA members, onsite in each participating school to supervise the tutors. The lesson learned was that relying on existing staff within DCPS to provide training to DC Reads tutors and others did not allow the program to deliver the training required in a timely fashion.

5. Frequent and Regular Tutoring Sessions

Establishing a specific time and days for the tutoring sessions was an effective practice in CIS' DC Reads program. The lessons learned related to the actual number of tutoring sessions that were held, as it was lower than expected. The actual length of each tutoring session was also less than the specified amount of time and the length of sessions varied from school to school, with some schools scheduling 40-minute and others 60-minute sessions.

6. Careful Evaluation, Assessment, Monitoring, and Reinforcement of Progress

The effective practices related to evaluation, assessment, monitoring, and reinforcement of progress were establishing uniform procedures and forms to document student progress and using student SAT 9 scores as a basis for teacher referral to the program. Another effective practice was making needed program modifications at mid-year—implementing the blended curriculum. One lesson learned related to the overall monitoring of tutors and program quality as this was not yet performed to the extent program managers believe is necessary. A second lesson learned was related to the program evaluation. An important outcome evaluation finding was that the group-administered SAT 9 tests provided useful information about student progress.

7. Access to Training and Technical Assistance Resources

In the CIS DC Reads program, one effective practice related to training and technical assistance is within-school access to a reading specialist. Another is access to the program's trainer reading specialist for additional support. The lesson learned was that while in theory a school-based reading specialist was available to the program in most schools, the reading specialist was not actually involved in many schools.

8. Engendering Positive, Caring Relationships Among Students, Staff, and Tutors

Tutors worked to develop positive relationships with the students they tutored. School staff and others believed that there were a number of academic and other benefits to the children from having the opportunity to interact one-on-one with a caring young adult. The use of Federal Work-Study students as tutors was an effective practice. The lessons learned were that tutor recruitment was difficult and that tutor turnover did not allow established relationships to flourish.

D. Conclusion

Students who received tutoring made significant gains in reading, and overall the program experienced considerable success in its second year. The experiences of the partners implementing the DC Reads program offer valuable lessons to others involved in planning or implementing similar programs. In Year 3, the partners expect to continue to build on previous successes while addressing the remaining challenges, primarily in the areas of tutor recruitment and retention and communication among partners.

Chapter I. Student Outcomes

A. Outcome Study Overview

The Year 2 evaluation of DC Reads focused on programs operated in 16 targeted assistance schools by Communities in Schools (CIS). Four of the schools served as primary study sites for the evaluation.

Students who participated in DC Reads were those identified by school staff as having low reading skills and being most in need of tutoring.

The Year 2 evaluation concentrated on assessing student outcomes[1], although it also included a program implementation assessment and an investigation of program achievements.

The purpose of the outcome assessment was to determine whether tutored students' reading skills improved after participating in the DC Reads tutoring program. The central element of the outcome assessment was the Reading Performance Battery.

Macro International Inc. conducted the Year 2 evaluation of DC Reads for the Corporation for National Service. The evaluation focused on programs operated by Communities in Schools (CIS). CIS operates DC Reads programs in 16 schools designated by DCPS in 1997 as the schools in most need of academic assistance. Four of the 16 schools served as the primary study sites for the evaluation. CIS was also the focus of the Macro's Year 1 evaluation of DC Reads.

Tutored students were identified by principals or teachers as those being in the greatest need of assistance in reading. Identification was based upon students' previous year Stanford Achievement Test (SAT 9) scores, below average reading grades in the previous year, poor in-class reading performance, home environment, and other factors.

In each of the four study schools, Macro randomly sampled 20 tutored students to participate in the outcome study. On the basis of fall 1998 SAT 9 scores, tutored students who participated in the outcome study did not differ significantly in reading achievement from those who did not. However, students selected to receive tutoring did have substantially lower reading skills than nontutored students.

---

[1] Measures of student outcomes included individually administered tests in a Reading Performance Battery and the Stanford Achievement Test (SAT 9). The tests in the Reading Performance Battery were administered by qualified educational testing specialists specifically for the evaluation. The SAT 9 tests are routinely administered to all students in the DC Public Schools in the fall and spring of each school year. Other data related to student outcomes included an assessment of students' attitudes toward reading, students' reading grades, and school attendance records.

BEST COPY AVAILABLE

B. Data Sources for the Outcome Study

1. Reading Performance Battery

The battery of tests used to assess reading performance consisted of eight standardized assessments. The tests were individually administered by professional educators to students at the four schools selected as primary study sites. Most of the tests were administered three times during the school year: Wave 1 in October and November 1998 (n=85), Wave 2 in January and February 1999 (n=75), and Wave 3 in May 1999 (n=51). The timing of testing and tutoring was as follows:

Wave 1 testing — Begin tutoring — Wave 2 testing — End of tutoring — Wave 3 testing

Figure 1: Reading Performance Battery Tests

Identification of Capital Letters

This test measures whether the student can recognize the 26 capital letters. Students are asked to identify randomly ordered capital letters. Scores are expressed as the percentage correct.

Identification of Lower Case Letters

This test measures whether the student can recognize the 26 lower case letters. Students are asked to identify randomly ordered lower case letters. Scores are expressed as the percentage correct.

Yopp-Singer Test of Phoneme Segmentation

This test provides a tool for assessing children's phonemic awareness and identifying children who may experience difficulty in reading and spelling. The test measures a child's ability to articulate the sounds of a spoken word separately and in order. For example, given the orally presented word "sat," the child should respond with three separate sounds: /s/-/a/-/t/. (The sounds, not the letter names, are the appropriate response.) There are 22 items in the test. Administration of the test normally requires about 5 to 10 minutes per child. Scores are expressed as the percentage correct.

Woodcock Reading Mastery Test Subtests (Forms G and H)

Various types of scores are available for all subtests of the Woodcock Reading Mastery Test, including standard scores and grade equivalent scores. For most analyses in this study, standard scores were used. The following subtests were administered:

Word Identification. The Word Identification subtest requires the student to read isolated words. The items range in difficulty from words typically presented in beginning reading programs to words difficult even for above-average college students. In Form G, there are 106 words on the list, beginning with very common monosyllabic words and ending with infrequently used multisyllabic words.

[2] Most tests in the Reading Performance Battery were administered three times during the school year. The exceptions are the Oral Reading Time and Comprehension subtests of the Durrell Analysis of Reading Difficulty. This test, available in only one form, was administered only in Wave 1 and Wave 3.

BEST COPY AVAILABLE

Word Attack. The Word Attack subtest measures students' ability to apply phonic and structural analysis skills in pronouncing words not recognizable by sight. The test includes both nonsense words and words that are used very infrequently. The word list begins with words that require only basic decoding skills and progresses to words that are much more difficult to decode.

Passage Comprehension. The Passage Comprehension subtest measures the student's ability to study a short passage – usually two or three sentences long – and to identify a key word missing from the passage. About one-third of the Passage Comprehension items are one sentence long and contain a picture related to the text. Forms G and H of Passage Comprehension each contain 68 items, arranged in order of difficulty.

Durrell Analysis of Reading Difficulty Subtests

Two subtests of the Durrell were administered for the study.

Oral Reading Time

For this test, students read prose aloud and then answer corresponding comprehension questions. The test presents eight passages in order of increasing difficulty. Students continue to read passages and answer questions until they make seven or more errors, or until the time required for reading the paragraph is outside the time norm. The instructional reading level is defined as the median grade on the time norm. The levels are, Primer, 1L, 1M, 1H, 2L, 2M, 2H, 3L, 3M, 3H, and so on. (In each case L stands for Low, M for Medium, and H for High, so that 3L indicates a low third-grade level). The scoring is determined by taking into account factors such as reading rate and number of word errors.

Oral Reading Comprehension

This test uses the same passages as Oral Reading Time. Students are assigned a score of "good," "fair," or "poor." If there are no more than two errors in answering the questions, the student's comprehension is considered "good." If three questions are unanswered or answered incorrectly, the student's comprehension is considered "fair." More than three questions unanswered or answered incorrectly results in a rating of "poor."

2. Stanford Achievement Test Scores and Language Arts Grades

We obtained the Stanford Achievement Test (SAT 9) reading scores, as well as the language arts grades and attendance for all second graders in the four study schools. The SAT 9 is used by DC Public Schools (DCPS) to assess its overall performance. A primary reason for obtaining the SAT 9 scores was to determine whether there was a strong correlation between the SAT 9 subtests and individually administered tests.[3] If so, this would have considerable cost implications for future evaluations of DC Reads and other tutoring programs. We also obtained students' grades in language arts in order to ascertain whether there is a strong relationship between students' reading performance (as measured by the Reading Performance Battery) and their grades.[4]

---

[3] Tests selected for the battery were individually administered to ensure that the results would not be affected by how well a child could follow directions in a group setting. A strong correlation between the SAT 9 and the battery would indicate that students had sufficient test taking skills and that both sets of tests measure the same constructs.

[4] In addition to students' grades, the attendance records of all second grade students were obtained. The school attendance levels of students who were tutored were not significantly different from the school attendance levels of their generally higher performing nontutored classmates. We also assessed tutored students' attitudes toward reading, but found no statistically significant change from program start to program end. There was also no systematic relationship between students' attitudes toward reading and their reading performance.

BEST COPY AVAILABLE

C. Changes on the Reading Performance Battery: Comparison of Wave 1 and Wave 3

Between Wave 1 and Wave 3, tutored students demonstrated statistically significant improvement on five of the eight measures in the Reading Performance Battery. Students improved in letter identification, phonemic awareness, word attack, and oral reading fluency and comprehension. Their most dramatic gains were in phonemic awareness, which is considered to be a fundamental skill, prerequisite to the development of more advanced reading skills.

There were no statistically significant declines in scores, even on the standard scores of the Woodcock subtests, which are adjusted for time on a monthly basis. On six of the eight tests, the standard deviation of scores decreased, a change indicating that students performed more consistently as a group in Wave 3 testing.

1. Summary of Reading Performance Battery Test Scores

Table 1 summarizes the scores for all three waves of testing for the Reading Performance Battery tests that have scores expressed as percentages or standard scores[5]. (Results for the Durrell Oral Reading Time and Oral Reading Comprehension are presented in the following section.)

Table 1: Reading Performance Battery–Means[6] for Wave 1, Wave 2 and Wave 3

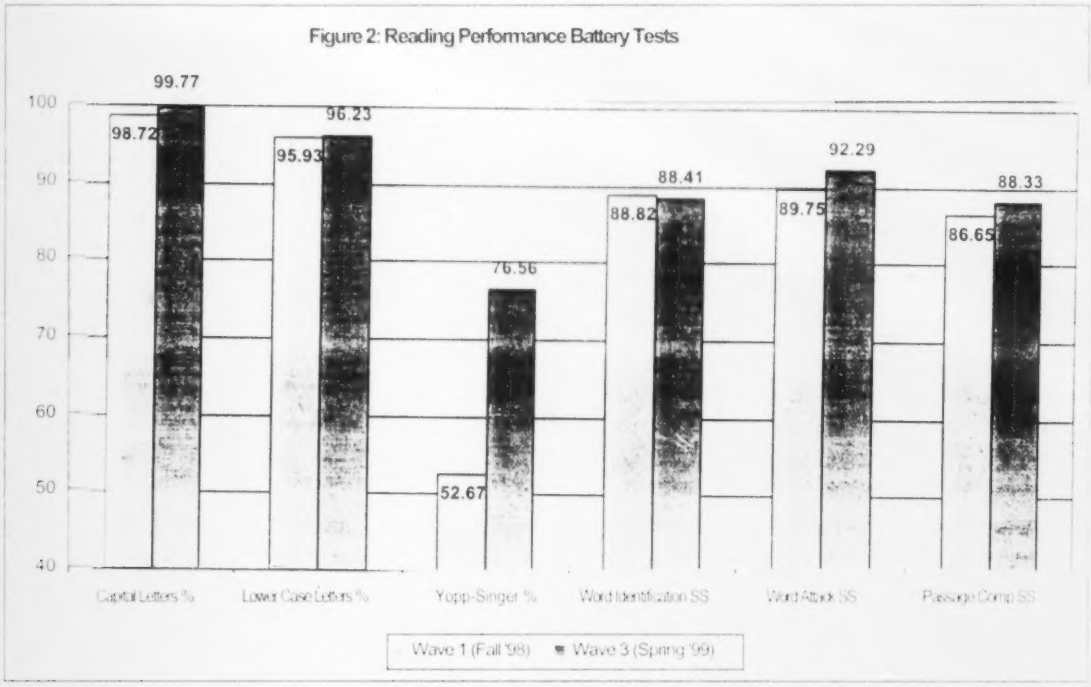
Test	Wave 1	Wave 2	Wave 3
Percentage of Capital Letters	98.72	98.49	99.77
Percentage of Lower Case Letters	95.93	96.38	96.23
Percentage of Yopp-Singer items	52.67	72.01	76.56
Word Identification Standard Score	88.82	89.82	88.41
Word Attack Standard Score	89.75	91.29	92.29
Passage Comprehension Standard Score	86.65	86.49	88.33

Figure 1 presents the comparison between Wave 1 and Wave 3 for the six tests included in Table 1.

[5] The Woodcock Reading Mastery Test yields many types of scores for its subtests, among them the Word Identification, Word Attack, and Passage Comprehension subtests used in this study. Standard scores have a mean of 100 and a standard deviation of 15.

[6] The mean is the average of scores or other numerical data (e.g., grades, days absent). It is obtained by adding up the values and dividing by the number of cases.

BEST COPY AVAILABLE



Grade Equivalent Scores for the Woodcock Reading Mastery Test subtests are summarized in Table 2 below. Overall, for the three Woodcock subtests, second-grade, tutored, tested students changed from 1.60 level (first grade, sixth month) before the tutoring program to a 2.17 level (second grade, second month). This means that students on average gained approximately 6 months in grade equivalency during their 6 months of participation in the DC Reads tutoring program and attendance in DC public schools. This is the amount of progress expected for an average student, but is noteworthy for the low-performing students served by DC Reads. On average, during the time in which they were tutored, students did not lose ground, as is often the case with low-performing students.

Table 2: Woodcock Reading Mastery Tests–Mean Grade Equivalent Scores

Test	Wave 1	Wave 2	Wave 3
Word Identification	1.74	2.03	2.20
Word Attack	1.68	2.06	2.36
Passage Comprehension	1.42	1.60	1.95

BEST COPY AVAILABLE

2. Significant Gains Between Wave 1 and Wave 3

Students demonstrated significant gains on five of the eight tests in the Reading Performance Battery between fall (Wave 1) and spring (Wave 3). We begin by discussing the results for the three tests whose scores can be represented as a percentage or standard score.

Identification of Capital Letters. In both Wave 1 and Wave 3 testing, the tutored, tested students demonstrated knowledge of virtually all capital letters. Despite the ceiling effect[7] demonstrated in Wave I, 11 out of 51 students improved on Wave 3 testing and the spread of students' scores narrowed greatly. Means scores are 98.72 percent for Wave 1 and 99.77 percent for Wave 3.

Yopp-Singer Test of Phoneme Segmentation. The tutored, tested students improved greatly in phonemic awareness, with the mean score increasing from 52.67 percent in Wave 1 to 76.56 percent in Wave 3. Most students (40 out of 51) improved their phonemic awareness from Wave 1 to Wave 3 testing.

Word Attack. Students' standard scores on the Word Attack subtest of the Woodcock Reading Mastery Test also increased. Mean scores increased from 89.75 in Wave 1 to 92.29 in Wave 3.

Table 3 presents the mean scores for Wave 1 and Wave 3, and shows the number and percentage of students who improved on each test.

Table 3: Change in Tutored, Tested Students from Wave 1 to Wave 3

Test	Testing Wave	Total Mean	Changes Mean	N Improved	% Improved
Percent of Capital Letter	Wave 1	**98.72**			
	Wave 3	**99.77**	**1.06**	11	22%
Percent of Lower Case Letter	Wave 1	95.93			
	Wave 3	96.23	0.30	18	35%
Percent of Yopp-Singer	Wave 1	**52.67**			
	Wave 3	**76.56**	**23.89**	40	78%
Word Identification	Wave 1	88.82			
	Wave 3	88.41	-0.41	18	35%
Word Attack	Wave 1	**89.75**			
	Wave 3	**92.29**	**2.55**	32	63%
Passage Comprehension	Wave 1	86.65			
	Wave 3	88.33	1.69	26	51%

(Statistically significant differences are in **boldface**.)

[7] When students score very highly on a measure, such as is the case with letter identification, there is very little room for improvement. This is often referred to as the "ceiling effect."

We found it relatively easy to understand why students did not show statistically significant improvement on Identification of Lower Case Letter and on the Passage Comprehension subtest of the Woodcock Reading Mastery Test. It is likely that the ceiling effect is a major factor in the very small and insignificant improvement in lower case letter identification. Regarding the other test, as described previously, comprehension must be preceded by the ability to decode with some facility, and many students in the study were still struggling with decoding.

Students improved greatly on the Yopp-Singer Test of Phoneme Segmentation. This test assesses whether children can segment the individual sounds of phonemes. They also improved significantly on the Woodcock Word Attack subtest, which measures how well students blend phonemes. However, they did not improve on the Woodcock Word Identification subtest. This test requires identification and blending of phonemes in words rather than in single syllables and is more complex. The Yopp-Singer and Word Attack tests require children to decode words that require a similar level of skill to "crack the code." All of the words on the Yopp-Singer and all but two of the words on the Word Attack test consist of a single syllable. The primary difference between the two tests is that the Word Attack subtest consists of nonsense words, which means that children will not be able to use their sight-reading vocabulary to read the words on the list. On the other hand, the Word Identification subtest has 14 two-syllable words and 4 three-syllable words, and many of them require more advanced decoding skills.

Students also demonstrated gains on two subtests of the Durrell Analysis of Reading Difficulty.

Oral Reading Time. Between Wave 1 and Wave 3, students improved significantly on this measure. Students became able to proceed further through the test, meeting the test's requirements for reading rate and accuracy. In terms of grade equivalents, students' median grade equivalent moved from first-grade middle (1M) to second-grade low (2L). This indicates that students gained approximately 6 months during their 6 months in the program and did not lose ground as often happens with low-performing students.

Oral Reading Comprehension. Students also improved significantly in their understanding of what they read aloud. Students ranged from poor to good in comprehension in Wave 1 testing, but all students in Wave 3 had good comprehension of what they read.

Table 4 presents the results for the Durrell Analysis of Reading Difficulty subtests.

Table 4: Change on Durrell Analysis of Reading Difficulty

Subtest	Wave	Mean[5]	Grade Level Median	Grade Level Range[6]
Durrell Reading Time	Wave 1	**3.10**	1M	1.6—2M
	Wave 3	**4.76**	2L	1.6—3M
Durrell Comprehension	Wave 1	**2.71**	Good	Poor—Good
	Wave 3	**3.00**	Good	Good—Good

(Statistically significant differences are in **boldface**.)

[5] The mean for Reading Time is expressed in minutes of reading aloud. The mean for Comprehension is based on the following assignment of values: 1=poor, 2=fair, 3=good.

[6] The lowest grade level score that can be assigned on this test is 1.6. A score of 1.6 means that the child is reading below first grade level. Possible scores within the range are 1.6, 1L, 1M, 1H, 2L, 2M, 2H, 3L, 3M.

Below are some further comments regarding changes from Wave 1 to Wave 3.

There were no statistically significant declines in scores, even in the standard scores of the Woodcock Reading Mastery Test subtests, which are adjusted on a monthly basis throughout the school year.

More than half of tested students showed improvement on tests of Phonemic Segmentation, Word Attack, and Passage Comprehension.

On six tests (Identification of Capital Letters, Phonemic Segmentation, Word Identification, Word Attack, Passage Comprehension, and Oral Reading Comprehension), the spread of students' scores decreased between Wave 1 and Wave 3. This indicates that tutored students performed more consistently as a group after being tutored and that fewer children were falling far behind.

D. Multivariate Analysis of the Effect of Tutoring on Reading Improvement

Macro employed Hierarchical Linear Modeling (HLM) for an analysis of the Reading Performance Battery tests that were administered in all three waves of testing. HLM is a statistical analysis technique that is useful when three or more waves of data are available. It was used to examine individual level growth in test scores, as affected by the amount of tutoring, while controlling for demographic and background variables of the student.

1. Overview of the HLM Approach

a. Rationale for Using Hierarchical Linear Modeling

A randomized experiment is the preferred method for attributing effects to tutoring since it provides, in the form of a control group, a basis for controlling for factors other than tutoring. However, it is not always possible to implement this type of design, as was the case for DC Reads. In such cases, regression analysis, which controls for background variables and other factors, is often used. HLM is a type of regression analysis, which is particularly suited to situations in which there are more than two waves of data collection. HLM allows researchers to examine and summarize differences across individuals using all three waves of information, which are represented through individual growth curves.

In this study, we tested students on three occasions (the minimum number required for HLM). One advantage of three or more data points is that it is possible to determine how students progress during various time periods. Another advantage is that it reduces the variation within individuals, that is, it increases reliability. As a result, it is possible to use a somewhat smaller sample than would be required for a study with only two waves of data collection.

b. Guiding Questions

The following questions guided our research approach.

1) How much on average do the reading skills of tutored children increase during the period of time they receive tutoring?

2) Are there differences among tutored students' growth rates?

3) What is the relationship between the initial reading skills and growth rate?

4) If there are differences in growth rates, then is growth rate related to DC Reads tutoring activities or to other factors such as attentiveness or gender?

c. HLM Models

We used a two-level HLM model in which the student's growth was modeled as a function of time. At the first level, we characterized each student's growth in reading skills in terms of the baseline score on the various tests, and the degree to which the student's score improved or decreased during the tutoring period on each of the inventories. We explored various representations of the growth curve, but generally used two specifications:

- A linear model that portrays reading improvements increasing steadily with time spent in tutoring sessions

- A quadratic model that tested whether the rate of growth increased as tutoring progressed.

At the second level, the individual growth parameters become the outcome measures, which may be related to some person-level demographic variables (e.g., gender) or to specific features of the tutoring program (e.g., the length of the tutoring program, writing, reading, review activities). This level 2 analysis explicitly examines variations in the growth curves across different students, and as such can provide information on what factors are associated with growth in reading skills.

2. Variation in Reading Skill Growth Rates

There was significant variation among tutored students in their initial scores on measures of phonemic awareness, word identification, and passage comprehension. For each of these outcome measures, students who had lower scores tended to grow at a faster rate than those with higher scores. This indicates that the lowest-performing students were "catching up" in some areas.

Using HLM we found that students grew at different rates on three of the six tests in the Reading Performance Battery administered in each of the three waves. (For the other three tests, there was either no evidence of growth and/or no variation in students' growth rates associated with their Wave 1 scores.)

Yopp-Singer Test of Phonemic Segmentation. There was significant improvement in students' phonemic segmentation scores between Wave 1 and Wave 3. The correlation between the Wave 1 status and growth rate was -0.935. This result means that students who had limited phonemic awareness at entry into the tutoring program tended to improve at a much faster rate than did students with better phonemic awareness.

Word Identification. In general, students experienced improvements in their word identification scores. The growth patterns indicate that, although there was a general improvement in word identification throughout the tutoring period, the rate of improvement decreased with time. Students who had limited skill in identifying words at entry into tutoring tended to gain on this measure at a faster growth rate (r=-0.776).

Passage Comprehension. There was significant variation on this measure in Wave 1 scores. Students who had lower passage comprehension at entry into the tutoring program tended to gain at a faster rate (r=-0.684).

3. Relationship of Specific Tutoring Activities to Reading Improvement

HLM provides some insight into which tutoring activities contributed most to reading improvement. In particular, students who practiced reading aloud more frequently during tutoring improved at a higher rate than students who did so less frequently.

Using HLM, we found a relationship between reading activity (as measured by the percentage of sessions in which students practiced reading aloud) and improvement on the standard score of the Passage Comprehension test. In addition, reading aloud was found to be associated with the quadratic terms of models relating to both Passage Comprehension and Word Identification. This relationship indicates that practicing reading contributed to individual level differences in growth toward the end of tutoring. Reading improved the ability of individuals to grow between the

second and third waves of testing. It should be noted that this effect, although significant, is small.

It should also be kept in mind that the lowest performers (those who worked in the Bob Book series) were the least likely to read continuous text aloud during their sessions. Thus, while these are important findings, we believe that they should be treated carefully since they account for a relatively small proportion of the total variation in growth rates.

E. Students' SAT 9 Reading Scores

Tutored students made greater gains on the SAT 9 reading tests than did nontutored students. On Normal Curve Equivalent scores (which compare students to all other students in the nation at the beginning and end of the school year), tutored students gained nearly twice as much as nontutored students. Tutored students gained nearly 7 points; the test's publisher considers an increase of 5 points significant. Also, 69 percent of tutored students improved on Total Reading, while just 60 percent of their nontutored peers demonstrated improvement.

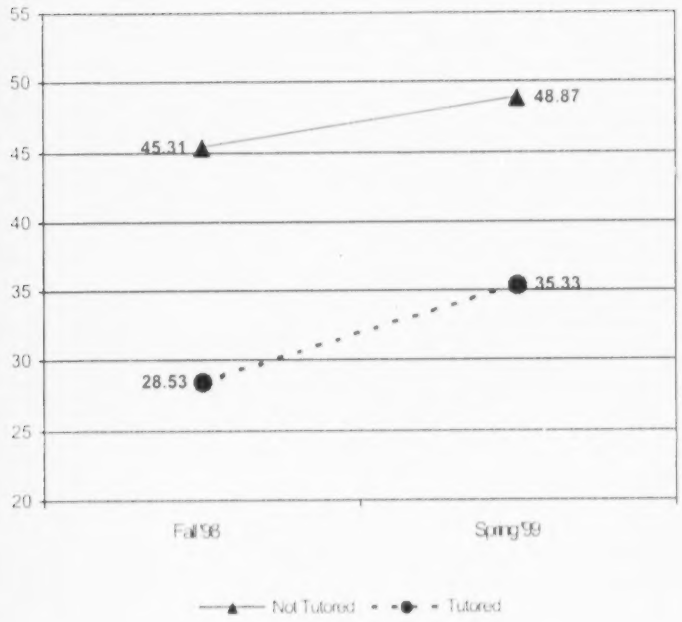
Although tutored students performed significantly below nontutored students in both fall and spring testing, their SAT 9 score improvements offer evidence that tutored students made significant gains above and beyond what would be attributable to reading instruction received in their classroom.

Mean NCE[10] scores of tutored students on Total Reading increased significantly, from 28.53 to 35.33. Both starting and ending scores for nontutored students were significantly higher, with these students going from 45.31 to 48.87. However, the score increase was nearly double for tutored students—6.80 for tutored students, as compared to 3.56 for nontutored students.

Because these scores compare students with others in the nation at the same time in the school year, this increase provides evidence that tutored students did gain some ground over the course of the school year. A representative of the Psychological Corporation of America, publisher of the SAT 9 tests, stated that it is expected that students will stay about the same from year to year on NCE scores. Anything over a 5-point gain is considered to be significant growth. The DC Reads program provided a 7-point gain, well beyond what is considered significant growth. Figure 3 depicts the mean score increases for tutored and nontutored students.

[10] Normal Curve Equivalent (NCE) scores have a mean of 50 and a standard deviation of 21.06.

Figure 3: NCE SAT 9 Total Reading Scores



To shed further light on the meaning of the SAT 9 score increases, Macro examined the scores of a subset of tutored children with a subset of nontutored children. We selected those students whose scores clustered around the Wave 1 mean for tutored students (i.e., around 28.53). For each group, we chose 20 children whose scores ranged from 23 to 34.4. Again, in this analysis, tutored students made greater gains—4.21 points, versus 2.31 for nontutored students. A caveat is that we do not know very much about the nontutored students. For example, we do not know if they had been identified as being in need of special education or if they were receiving tutoring from another source. However, in the absence of a control group, this comparison offers further evidence that tutored students did make gains beyond what would have occurred without DC Reads.

The literature on poor readers supports this view. One study (Fletcher & Satz, 1980)[11] found that, not only do poor readers not catch up, they fall further behind with each progressive school year. Nearly 40 percent of second grade children classified as "poor readers" on the basis of standardized tests became "severe" readers by fifth grade. Since that time, much research has indicated that children who are behind their peers in first grade tend to improve at the same rate as good readers, but that they never catch up (McGuinness, 1997)[12]. While there are more and

[11] Fletcher, J.M., & Satz, P. (1980). *Lag-deficit characterization of the disabled reader: Some alternative interpretations.* Paper presented at the 80th annual meeting of the International Neuropsychological Society, San Francisco, February.

[12] McGuinness, D. (1997). *Why Our Children Can't Read.* New York : The Free Press.

less pessimistic views of the prognosis for children who experience reading failure in the early elementary grades, virtually no one expects poor readers to become good readers without intervention. The SAT 9 score increases for tutored children provide evidence that they are "catching up" in some areas.

F. Students' Grades in Reading

> The language arts grades of tutored students increased somewhat, but the improvement was statistically significant only among the tutored students who were not individually tested by Macro educational testing specialists.[13]

The average grades in both reading and writing of all second grade students increased somewhat between the first and fourth quarters. The increase was statistically significant for nontutored students and for tutored students who were not tested by Macro. Despite the gains for tutored students, even at the end of the school year, nontutored students were still significantly outperforming tutored students. (Differences in grades between tutored and nontutored students are significant at the p=.00 level in both the first and fourth quarters.) On average, the grades of tutored students indicate that they were "approaching the standard" for the second-grade level at the end of the school year.[14]

Table 5: Comparison of Grades of Tutored and Nontutored Students

Student Classification	Reading Grades		Writing Grades	
	First Quarter	Fourth Quarter	First Quarter	Fourth Quarter
Tutored (N=76)	1.82	1.96	1.88	1.97
Not Tutored (N=84)	2.59	2.71	2.57	2.65

G. Summary of the Outcome Study

Tutored students demonstrated gains on a number of measures used in the outcome study, and there is evidence that they were "catching up" in some areas.

> Tutored students made significant gains on five of the eight measures in the Reading Performance Battery.

---

[13] Macro randomly sampled 80 second-grade students (20 in each of four schools serving as primary study sites) to be individually tested for the outcome study. There were also other second-grade students who received DC Reads tutoring in the four schools. While first quarter SAT 9 test scores indicate that the two groups were comparable in reading skills, the grades of those who were individually tested did not increase as much as those who were not.

[14] Integer numerical grades ranging from 1 to 4 are used in DCPS elementary schools. 1 = Child is not meeting the standard for this grade level; 2 = Child is approaching the standard; 3 = Child is meeting the standard; and 4 = Child exceeds the standard.

The standard deviation decreased on six of the eight measures in the battery. This means that the spread of students' scores decreased, that students were performing more consistently as a group at the end of tutoring, and that fewer students were falling far behind.

On three of the six tests used in the HLM analysis, tutored students who entered with low skills improved at a faster rate than tutored students who entered with higher skills.

The percentage of sessions that included reading aloud was associated with improvement on two of the five measures included in the HLM analysis and, in particular, appears to have contributed to growth between the second and third waves of testing.

The SAT 9 scores of tutored students increased significantly and at nearly twice the rate of nontutored students.

The language arts grades of some tutored students improved significantly from the first to fourth quarter.

Chapter II. Program Description

DC Reads is a partnership between the Corporation for National Service, the District of Columbia Public Schools, Communities in Schools, and eight local universities.

DC Reads tutors, most of whom were Federal Work-Study students attending partner universities, provided tutoring to early elementary students twice a week. Tutors used a curriculum and instructional materials that the partners purchased and adapted.

At the beginning of the 1998-99 school year, two study schools used the Book Partners curriculum and two used the Reading One-to-One curriculum. At the beginning of the second semester, a new blended curriculum was adopted in all schools. The blended curriculum enabled tutors to meet the needs of nearly all students who received tutoring, including most of those with very low reading skills.

A. The Partners

The DC Reads program was first implemented during the 1997–1998 school year as a partnership between the:

Corporation for National Service

District of Columbia Public Schools (DCPS)

Communities in Schools (CIS), a nonprofit organization

Six local universities, with Georgetown University acting as lead.

Other partner universities were American, Catholic, George Washington, Howard, and Trinity. During the 1998–1999 school year, the program was expanded to include two more city-based universities, Southeastern and the University of the District of Columbia. It also expanded to include more nonprofit community organizations that coordinate the delivery of tutoring services in additional DC public schools, as well as in charter and private schools.

B. Program Operations

The DC Reads program is intended to address the needs of first-, second-, and third-grade students who have low reading skills, but do not have a disability that impedes reading performance. During the 1998-1999 school year, there was a strong focus on providing tutoring to second-grade students who are the focus of the Year 2 evaluation. Students who were served by DC Reads in Year 2 were to receive tutoring twice a week, in their schools, in space assigned for the operation of the DC Reads program. CIS coordinated the program and provided supervision to Americorps*VISTA members through project site directors. AmeriCorps*VISTA members provided onsite supervision to tutors, most of whom were Federal Work-Study students

attending partner universities, some of whom were community volunteers. Tutors were typically at their assigned schools for 2 hours; in that time, they were to provide tutoring to two or three children. Tutors used a prescribed tutoring curriculum and materials that were purchased and adapted by the partners. Tutors received training in the use of the instructional materials and in other topics related to tutoring in reading from the CIS training coordinator.

Overall, Macro estimates that DC Reads provided tutoring to approximately 504 students in its 16 CIS school sites during the 1998–1999 school year. A total of 340 tutors provided tutoring to those students. In the four study schools, 136 students received tutoring from 74 tutors.

C. How Children Learn To Read

Learning to read in English is a very complex process. A first step toward becoming a proficient reader is to develop an awareness of the 43 English-language phonemes (the basic units of speech) and the ability to segment and blend phonemes in words. Phonemic awareness involves acquiring knowledge both of the basic sounds and the letter symbols that represent them. Typically children are focused on phonemic awareness at the very beginning of learning to read and, in their reading instruction, work with very simple real and nonsense words such "cat" and "mop." However, English has many sounds that are not predictable. Children need to learn a basic code of reading, which includes digraphs, redundant consonants, irregular plurals, and diphthongs.

In the early grades, in order to decode effectively, children have to develop an understanding of the probability structure of English spelling. After they have gained a knowledge of the basic code of English spelling, children must learn to decode multisyllable words and to readily recognize those relatively few true "sight words" (e.g., friend, leopard) for which none of the decoding rules applies. Decoding English is a challenge, and yet it is something that all children must learn to do, because accuracy and speed in decoding are prerequisites to reading comprehension. In fact, research indicates that the best predictor of children's reading comprehension on standardized reading comprehension tests is their ability to decode individual words out of context.

The range of materials used in CIS school sites accommodates the needs of a wide range of children. Figure II-1 in Section E shows the match between reading development stage and specific curriculum materials. Section D, immediately following, describes the materials used.

D. Curriculum and Instructional Materials

In Year 1, all CIS DC Reads schools used the Book Partners curriculum, which was an adaptation of the Book Buddies program developed at the University of Virginia. At the beginning of Year 2, a decision was made to test a new curriculum in two schools. The new curriculum was based on Reading One-to-One, which had been developed by George Farkas at the University of Texas.

Following is a brief description of the two curricula and the materials that were used in tutoring sessions.

1. Book Partners

Each 60-minute tutoring session in the Book Partners curriculum included three components:

1) Rereading familiar material

2) Early Success reading

3) Choice Time reading.

The second component, Early Success reading, included the following activities to build literacy skills: coached reading, shared reading, independent reading, working with words, and writing. The primary materials in the Book Partners curriculum are:

Early Success Books. A series of 30 books specifically designed for use in primary reading instruction.

Choice Books. A collection of books of authentic children's literature, many of which have been recognized for excellence.

All children referred for tutoring began with Book 1 of the Early Success series. In most cases, once a child mastered an Early Success book, the child moved on to the next book in the series. For some children who had good basic decoding skills, books were skipped and additional time devoted to Choice Books. In all cases, after completing the Early Success lesson for the day, the child could select a Choice Book to read with the tutor.

2. Reading One-to-One

A typical 40-minute session in the Reading One-to-One curriculum included:

1) Re-reading familiar books

2) Reading a new book

3) Drill of letters, sounds, and words

4) Writing activity.

There are three levels through which students could progress: Alphabet, Word Family, and Reading Comprehension. Each level has an associated set of books.

Bob Books. A series of 12 books intended for students who have difficulties identifying letters and their associated sounds.

Steck-Vaughn Books. A series of 40 books. Each book has its own sound word pattern focus, which includes both single sounds and blends.

Reading Comprehension Books. A series of books, known as the Sunshine Series, which was developed for use in reading improvement programs and was informally divided into levels by its publisher.

Upon entering the program, each child was tested and placed into one of the three levels. Most of the children participating in DC Reads were initially placed in either the Bob books or the Steck-Vaughn books. Students' reading skills were assessed every fifth session and students advanced to a higher curriculum by attaining high scores on the assessments.

3. Other Instructional Materials

For both curricula, in addition to the books described above, tutors used other materials that were prescribed in the lesson plan associated with each book. The materials included notebooks, word lists, letter tiles trays, sound boxes, bingo cards, sentence strips, and flashcards.

4. The Blended Curriculum

In January 1999, a blended curriculum combining both Book Partners and Reading One-to-One was adopted in all 16 schools. This change was made because the DC Reads partners decided that neither curriculum was meeting the needs of all children and that the resources of the two curricula should be combined. In particular, some students had very low reading skills and needed to start with very basic materials. The Reading One-to-One curriculum included such materials. The new combined curriculum used both the materials and instructional techniques of Book Partners with the materials and regular assessments of the Reading One-to-One curriculum.

E. Session Activities

On the basis of our analysis of the progress logs for students involved in the outcome study, tutors used the Bob Books least frequently (48 sessions) and the Steck-Vaughn books most frequently (574 sessions). Tutors used the Early Success books in 383 sessions and the Reading Comprehension books in 206 sessions.

In terms of addressing the needs of children at various levels, Figure II-1 displays the relationship between the primary skill development focus and instructional materials, and the measures of student reading performance presented in Chapter I.

Figure II-1: Relationship of Reading Stage, Curriculum Materials, and Outcome Measures

Reading Development Stage	Curriculum Materials	Outcome Measure
Phonemic Awareness →	Bob Books Steck-Vaughn	Yopp-Singer Test of Phonemic Segmentation Woodcock Word Attack
Basic Decoding →	Steck-Vaughn Early Success	Woodcock Word Attack Durrell Oral Reading Time SAT 9 Word Study
Advanced Decoding →	Early Success Choice Books	Woodcock Word Identification Durrell Oral Reading Time SAT 9 Word Reading
Comprehension →	Choice Books Sunshine Reading Series	Woodcock Passage Comprehension Durrell Oral Reading Comprehension SAT 9 Reading Comprehension

Overall, tutored students received about 20 sessions (approximately 15 hours) over a 6-month period. This is about one session per week on average. Sessions lasted an average of 45 minutes. The most common activity in the sessions was reading aloud from a lesson or choice book. For a child who received 20 sessions, typically 18 of 20 would include reading[1], 17 of 20 would include working with words, 17 of 20 would include writing, and 16 of 20 would include review activities.

Students were generally quite attentive during tutoring sessions. Tutors were asked to rate the child's attentiveness at the conclusion of each session. Across all sessions the average attentiveness score was 2.51 on a 3-point scale (1=poor, 2=average, 3=excellent).

[1] Reading is defined as reading aloud to the tutor from a book in one of the book series (Early Success, Bob, Steck-Vaughn, and Sunshine Reading Comprehension) or from a Choice Book. Students at lower levels did this less often. Just 65 percent of lessons with the Bob Books included reading aloud. For the other books series, the percentages increase dramatically—92 percent for Early Success, 94 percent for Steck-Vaughn, and 99 percent for Sunshine Reading Comprehension.

Chapter III. Effective Practices and Lessons Learned

This chapter describes the various accomplishments and challenges of the CIS DC Reads program during its second year of operation in 1998-1999. To better assist readers in understanding these accomplishments and challenges, the information is presented using the Corporation's Effective Practices Model, which describes effective practices of the program, and the various lessons learned. Information for this chapter is drawn from the implementation assessment and investigation of program achievements.

A. Overview of the Implementation Assessment and Investigation of Program Achievements

Along with the outcome study, Macro conducted an assessment of CIS' implementation of DC Reads during the 1998-1999 school year and documented the program's achievements.

Implementation Assessment. The purpose of the implementation assessment was to provide 1) information for program improvement, 2) a context for interpreting the results of the achievement and outcome data, and 3) lessons learned for other programs that provide tutoring in reading.

Investigation of Program Achievements. The evaluation team examined the achievements of the DC Reads program in terms of number of students tutored, tutors recruited, and hours of tutoring provided. They also surveyed those associated with the program to obtain their perceptions about the achievements of DC Reads.

Macro collected data related to program implementation and achievements by means of interviews, observations of tutoring sessions, various program recordkeeping logs, and survey questionnaires. The principals, teachers, AmeriCorps*VISTA members, tutors, parents, campus coordinators, and project site directors associated with the four primary study sites were surveyed[1]. A full discussion of the study design is presented in Appendix A: Methodology. Most of the instruments that were used are included in Appendix C: Instruments. The standardized reading tests that were used to assess students' reading skills are not included.

The remainder of this chapter is devoted to describing the program's effective practices and reporting the lessons learned through the implementation assessment and investigation of program achievements.

[1] The response rates for teachers, tutors, and parents were below 50 percent, and it is quite possible that the results may have been biased because of this. Although the people who responded were generally quite positive about the program, nonrespondents may differ.

B. The Effective Practices Model

The Corporation previously identified eight principles of high-quality tutoring programs in its *Principles and Key Components for High-Quality America Reads and National Service Program Initiatives*. These principles are grounded in the literature on reading research and are shown in Figure 4.

Figure 4
Eight Principles of High-Quality Tutoring Programs

1) Use of research-based elements to produce reading achievement
2) Well-structured tutoring sessions in which the content and delivery of instruction are carefully planned
3) Close coordination with schools, school administration, and the classroom and/or reading teacher
4) Intensive and ongoing training and supervision for tutors
5) Frequent and regular tutoring sessions
6) Careful evaluation, assessment, monitoring, and reinforcement of progress
7) Access to training and technical assistance resources
8) Engendering positive, caring relationships among students, staff, and tutors

The Corporation defines ***effective practices*** as the activities and methods a program engages in to produce positive results. Practices can be divided into two types, those that describe the program model and those that describe the delivery system.

> The program model (the *what* of the program) includes the features and activities that are directly related to the benefits and outcomes of the program and thus those that are directed toward the targeted service recipients – in this case, primarily the children who are tutored.
>
> The delivery system practices (the *how* of the program) are descriptors that relate to the program implementation, including organizational infrastructure, and the features and activities that relate to assessment and evaluation.

In this instance, the program model is the CIS DC Reads program, which used both the Book Partners and Reading One-to-One tutoring curricula along with the various reading, writing, and comprehension activities. The delivery system that made the tutoring possible included a collaborative organizational structure that brought the talents of community services

organizations, universities, public schools, AmeriCorps*VISTA members, and tutors to bear on the common purpose of improving the reading performance of academically and socially disadvantaged second-grade children. The assessment included tests of students' reading abilities at study sites, as well as observation data, surveys, interviews, and logs from a range of program stakeholders.

Lessons learned are the solutions implemented in the program or envisioned as improvements for future programs, often developed in response to challenges that the program confronted. Lessons learned may include solutions that were actually put into practice and those that are planned, as well as insights into aspects of the program that were not as successful as anticipated.

Appendix B contains a table that lists the sources of information that were used to document each of the effective practices and lessons learned presented in this chapter.

C. Eight Key Components for Reading Tutoring Programs

1. Use of Research-based Elements to Produce Reading Achievement

The CIS DC Reads program used two tested reading curricula to improve the reading abilities of low-performing second-grade students. The selection and use of well-tested curricula comprised the effective practice component of the program; the adaptation of the model to better meet the needs of students reflects the lessons learned.

a. Effective Practices

During the first year of the CIS DC Reads program, only one curriculum, Book Partners, was used. A second curriculum—Reading One-to-One—was introduced into the program at the beginning of the 1998-1999 school year (the second year of the program). Both curricula had been tested in other states and were in use in other school districts. These curricula included designated sets of books for use on specific tutoring activities.

Effective Practices

✓ Use of well-tested curricula developed at major universities

Lessons Learned

✓ Blending two curricula to better meet student needs

✓ May still require additional materials to meet the needs of some children

Book Partners. Book Partners was originally adapted from the Book Buddies program developed at the University of Virginia. Each Book Partners tutoring session consists of three components: rereading familiar material, reading from an Early Success book, and Choice Time reading. The Early Success portion of the tutoring session includes coached reading, shared reading, independent reading, working with words, and writing. The books included as part of the curriculum were briefly described in Chapter 2.

Reading One-to-One. Reading One-to-One was developed by Dr. George Farkas of the University of Texas at Dallas in 1991 and has been used in Texas since that time. Most tutoring sessions included several activities, including rereading familiar books; reading a

new book; drill of letters, sounds, and words; and a writing activity. The curriculum required tutors to focus on key elements in the reading process: learning letter/sound relationships, using sound sense to read words in connected text, and understanding what is read. Students move through the curriculum by progressing through three "levels"—alphabet, word families, and reading comprehension. Students were able to advance once they demonstrated their readiness through an assessment. The curriculum included skill assessments at periodic intervals. The books included in the curriculum were described in Chapter 2.

b. Lessons Learned

The second year of the DC Reads program offered two lessons learned. First, program partners and tutors found that the children served had diverse needs and a single curriculum did not meet the needs of all the students served by the program. Therefore, in late fall a decision was made to combine the resources of both curricula to create a merged curriculum. As a way of blending the curriculum, the Corporation encouraged both the introduction of the Reading One-to-One curriculum into the program at two schools as a test and subsequently expanded use of these materials at all 16 program sites. The decision to combine the resources was extremely well received by those who were directly involved with the children receiving tutoring.

Despite the merging of the two tested curricula, some children continued to have reading needs that were not fully addressed by the program. One group of children referred to the program had more severe reading difficulties than the program had anticipated. Another group of children were those who progressed but did not progress enough to be graduated from the program.

Although no one program can meet the needs of all children, these lessons learned indicate that further modifications to the curriculum (or eligibility criteria) may need to be made in subsequent years of the program.

2. Well-structured Tutoring Sessions in Which the Content and Delivery of Instruction Are Carefully Planned

The CIS DC Reads tutors provided reading activities to second-grade children using formal written lesson plans to guide each tutoring session. The curriculum materials were geared to a wide range of reading abilities; thus tutoring activities were generally at the appropriate instructional level.

a. Effective Practices

Both curricula had structured activities for tutors to follow. However, not all books within each of the two curricula had an available lesson plan. When the use of the Reading One-to-One books was extended to all DC Reads program sites, the reading specialist trainer developed lesson plans for each book used in the Reading One-to-One curriculum. As a result, for each book used in the merged curriculum, there was an associated lesson plan, which laid out all of the activities to be conducted in working with that book. The development and use of the lesson

plans made the program more consistent and not as dependent on the particular strengths of individual tutors.

The blended curriculum also provided a wider range of materials that tutors and children could use to work on specific reading skills. Another benefit of the blended curriculum was that it allowed the tutor to more closely match the reading activity to the skill level of the child. Thus tutoring sessions were likely to be at an appropriate instructional level (neither too hard nor too easy).

Effective Practices

- ✓ A curriculum that includes a wide range of materials to meet the varied needs of students
- ✓ A corresponding lesson plan for each book

Lessons Learned

- ✓ Developing individual lesson plans rather than using large curriculum manuals
- ✓ Recognizing the need to develop highly structured lessons to guide tutors

b. Lessons Learned

At the beginning of Year 2, DC Reads program staff had already developed lesson plans for all of the Early Success books used in Book Partners schools. It was expected that those schools using the Reading One-to-One curriculum would use the tutor manual developed for that program. However, tutors became frustrated trying to use the manual to guide their sessions and found it unwieldy. As a result, when the curriculum was blended and the Reading One-to-One curriculum was extended to all DC Reads sites, the reading specialist/trainer decided to develop lesson plans for each book used in the Reading One-to-One curriculum.

The development and use of the lesson plans made the program more consistent than it had been in the first year. In the Year 1 evaluation, it was observed that tutors, at times, made up activities, in part because the range of materials was not sufficient to meet the needs of all children. Some tutors were more skilled at developing appropriate activities than others. In Year 2, with an increased range of materials and clear lesson plans, tutors could meet most students' needs by closely following the lesson plans.

3. Close Coordination With Schools, School Administration, and the Classroom and/or Reading Teacher

There were several effective practices associated with this program element. First, there was teacher involvement in selecting the students for the tutoring. Second, the DC public schools office was involved in the program, and encouraged principals to take an active role in supporting the program. The presence of AmeriCorps*VISTA members at each school provided a point person for school staff. In addition, provision of information to school staff at the beginning of the school year was also important.

Improving access to facilities and equipment, increasing communication with teachers, increasing access to the reading specialist, and recognizing how cultural differences impact communication comprise the lessons learned related to close coordination.

a. Effective Practices

Teachers were much more involved in the selection of students who received tutoring in Year 2 than they had been in Year 1. In the first year teachers had very little role in identifying the students who would participate. This was due in part to the attempt to use diagnostic testing scores and to randomly select students who would participate from those students who were in need of assistance but were not being served by other intervention programs. Because teachers objected to this approach and expressed the desire to have a say about which students were chosen, teachers were asked to nominate children to participate in DC Reads and to provide their reason(s) for the nomination in Year 2. The process of having teachers "refer" students to the DC Reads program strengthened the program and was heralded as an important improvement in the program.

Effective Practices

- ✓ Involving teachers in selecting students for tutoring
- ✓ Involving central school district offices in initial planning
- ✓ Actively involving school principals
- ✓ Having an on-site AmeriCorps*VISTA member as a coordinator at each school
- ✓ Providing information about the program at the beginning of the school year

The District of Columbia Public Schools office (DCPS) was a partner in DC Reads when the program began. The program was able to achieve implementation across the 16 program schools because of central office involvement and support. The central office ensured that all schools were able and willing to participate and would devote some resources to the program.

One of the single most important factors in the success of the program was active support at the principal level. The original involvement of the DCPS central office allowed program staff to build a solid relationship with each DC Reads school and its principal. Program staff have been able to maintain and enhance the initial relationships established with principals. Principals viewed the program very favorably and believed that it was a valuable addition to their schools.

The on-site presence of the AmeriCorps*VISTA members to coordinate DC Reads activities and to provide other support to the school was important to program success. The AmeriCorps*VISTA members were well regarded and accepted by school staff.

Although teachers and principals did not always remember the details of CIS' DC Reads program, most were able to recall learning about the program in early fall. The information was generally conveyed without a formal program orientation. The teachers' and principals' general awareness about the program fostered their support for the tutoring effort.

b. Lessons Learned

The program experienced some difficulties with facilities and access to equipment; the quality of tutoring space and consistency of space availability varied greatly from one school to another. With respect to the facilities, the program encountered problems such as the tutoring room being

over-heated, the program being "pre-empted" from tutoring/program space when other school needs arose, multiple children and tutors working in the same room with little separation of space, and the room being too small for tutoring activities. In addition, program staff had very limited supplies and had difficulty making needed copies or printing documents. In some schools, additional effort may be needed to arrange appropriate and consistently available space for the tutoring program.

Lessons Learned

- ✓ Improving access to facilities and equipment
- ✓ Increasing communication with teachers
- ✓ Increasing access to reading specialists
- ✓ Recognizing how cultural differences impact communication

A theme that emerged from the evaluation data was a desire for greater communication and collaboration with teachers in subsequent years. Teachers were concerned about coordinating the CIS DC Reads tutoring with classroom instruction and felt they needed additional information about the program curriculum. Tutors also felt there needed to be increased communication and collaboration with teachers. Although teachers and tutors were positive about the program and its effects, each group acknowledged the need to increase communication with the other.

Reading specialists were considered an important resource to the AmeriCorps*VISTA members and tutors. Despite the importance of these individuals in each school, the AmeriCorps*VISTA members and tutors felt more frequent contact and support than occurred during the school year is required, as not all reading specialists provided support to the program.

The differences in ethnicity/race and organizational cultures (e.g., universities, public schools, etc.) created challenges for program partners. During the second year of program operation, senior staff made efforts to decrease communication barriers by hiring more diverse staff and by providing a diversity training workshop. Efforts to recognize differences among individuals and organizations and to strengthen communication across groups will continue in Year 3.

4. Intensive and Ongoing Training and Supervision For Tutors

For this program element there were two effective practices. One effective practice was hiring an individual with a background in reading to coordinate all DC Reads training activities. Another effective practice was to have designated individuals, in this instance AmeriCorps*VISTA members, supervise the tutors. The lesson learned was that relying on existing staff within DCPS to provide training to DC Reads tutors and other staff, did not allow tutors and others to receive the level of training they required to achieve program success.

a. Effective Practices

Hiring a Reading Specialist for Training. In Year 2 of the DC Reads program, CIS created a new trainer/program development specialist position to fill the need to provide readily available training. The person selected to fill this position was a reading specialist who had experience working in the DC public schools. In addition to making important decisions about the curriculum, the trainer provided regular training to AmeriCorps*VISTA members and tutors.

The trainer established a training calendar and conducted four to five training sessions per week. The availability of a trainer and the frequency of training allowed new tutors to be trained immediately. Training sessions were held in various locations including onsite at participating schools and at a designated training facility at one of the elementary schools. Altogether, during the school year approximately 650 people received training.

Tutor training included an overview of child development, the reading skills development continuum, and specific tutoring strategies. AmeriCorps*VISTA members training topics included literacy, child development, discipline strategies, strategies for reluctant readers, the tutoring curriculum (Book Partners or Reading One-to-One), supervising tutors, team building, managing relationships, and diversity. There was an initial training at the beginning of the year and some follow-up training later in the year, including retraining tutors and AmeriCorps*VISTA members for use of the blended curriculum.

Effective Practices

- ✓ Establishing a training coordinator position
- ✓ Designating AmeriCorps*VISTA members as tutor supervisors

Lessons Learned

- ✓ Not relying on existing public school staff to provide timely training

Supervision of Tutors. One of the major responsibilities of the AmeriCorps*VISTA members was to supervise tutors. (In many instances, there were two AmeriCorps*VISTA members at a school.) In turn, the project site directors provided oversight and direction to the AmeriCorps*VISTA members working in CIS DC Reads schools. Typically, AmeriCorps*VISTA members met with tutors two to four times a month; in some instances, AmeriCorps*VISTA members met with tutors each time they came to the school to tutor. Project site directors, principals, teachers, and tutors generally believed that the supervision of tutors was satisfactory.

b. Lessons Learned

In Year 1 of DC Reads, there were initial centralized trainings for AmeriCorps*VISTA members and for tutors. Following that, DCPS staff provided training for AmeriCorps*VISTA members and tutors. These staff members had other responsibilities, and providing training to DC Reads AmeriCorps*VISTA members and tutors was not a high priority for them. As a result, training was not as accessible as CIS wanted it to be. It was particularly important to have training immediately available for new tutors so they could start in the program when their interest was high. By adding a reading specialist/training coordinator position at CIS, regular training became available, accessible, and targeted to the needs of the program.

5. Frequent and Regular Tutoring Sessions

For this program element, establishing a specific time and days for the tutoring sessions was an effective practice in CIS' DC Reads program. The lessons learned related to the actual number of tutoring sessions that were held, and the actual length of each tutoring session.

a. Effective Practices

The CIS DC Reads program had an established time for tutoring sessions. The tutors met with their students between the hours of 1 and 3 p.m. two days a week at participating schools. Teachers and principals supported this schedule because it allowed the students to participate in regular classroom reading activities, which were held in the morning. In most schools, the tutoring sessions fit in well with the school schedule. In all, more than 500 students received tutoring from nearly 350 tutors.

Effective Practices

- ✓ Establishing specific days and times for tutoring sessions

Lessons Learned

- ✓ Increasing the number of sessions in which tutored children participate
- ✓ Ensuring tutoring sessions are held for the designated length of time
- ✓ Reconciling tutoring session length between the two curricula

b. Lessons Learned

Overall, students participated in fewer tutoring sessions than had been planned. Students received approximately 20 45-minute tutoring sessions during the 1998-1999 school year or roughly 15 hours of reading tutoring. Several factors appear to have contributed to the lower-than-expected number of sessions. A major factor was difficulty with tutor recruitment and retention, which is discussed on page 13 in this chapter. Other factors included a typical mid-October start and late-April end for tutoring, school vacations and holidays (Federal holidays often occur on Mondays), special programs at the school, partial school days, and students' absences. In subsequent years, the tutoring calendar needs to take these factors into consideration in order to deliver more sessions.

Tutoring sessions typically did not last the full amount of time (40 or 60 minutes) designated for the tutoring session. This was because tutors were usually responsible for getting students from class and returning them to class at the end of the session. They also had to gather materials or get organized for the tutoring session. Options for addressing these issues include recruiting more tutors, offering more tutoring sessions per week, or assigning two tutors for every three children.

Each of the two curricula had a different scheduled session length for tutoring sessions. For the Book Partners curriculum, the tutoring sessions were scheduled to last for 60 minutes; for Reading One-to-One, the tutoring sessions were scheduled to last for 40 minutes. Each session length had advantages and disadvantages. The advantage of 40-minute sessions is that a tutor can conceivably tutor three students in a 2-hour tutoring block. The downside is that there is less time for casual interaction or mentoring or for Choice Time books. The tutor must focus on finishing the lesson. The 60-minute sessions allow the tutor more time to interact with the child, but a tutor can see only two children in the allotted 2-hour time block. Some tutors found the shorter sessions to be stressful and not as rewarding to them as the longer sessions. Program managers plan to standardize session length, having 50-minute sessions in all schools during Year 3.

6. Careful Evaluation, Assessment, Monitoring, and Reinforcement of Progress

The effective practices for this program element were establishing uniform procedures and forms to document student progress, using student Stanford Achievement Test (SAT 9) scores as a basis for referral to the program, and making needed program modifications mid-year. One lesson learned related to the overall monitoring of tutors and program quality. A second lesson learned was related to the program evaluation. An important outcome evaluation finding was that for second-grade children, the SAT 9 scores provided useful information. However, the specialized Yopp-Singer Test of Phonemic Segmentation (one of the Reading Performance Battery tests) showed that phonemic awareness was an area in which dramatic growth occurred. The magnitude of this growth was not detected by any of the SAT 9 tests.

a. Effective Practices

CIS managers, working with the program evaluation contractor, developed an easy-to-use Book Partners progress log form for use by the tutors. Although the Reading One-to-One curriculum came with its own set of progress monitoring and student assessment tools, once the two curricula were blended, tutors preferred using the newly developed log. The log provided space for tutors' notes about the student's progress and accommodated the varying instructional materials included in the blended curriculum. Each tutored child had a notebook in which progress logs were kept along with other information about the student's progress, such as assessments and word lists. Establishing an easy-to-use form to document progress helped improve the quality of documentation by tutors.

Teachers recommended students for individual reading tutoring using several criteria, among them low SAT 9 reading scores. Using these scores as a "screening" element for entrance into the program offered an objective measure of student need for the tutoring and decreased teacher reliance on more subjective measures, such as a student's home environment or grade in reading.

Effective programs collect information on a continuous basis to help program staff determine whether program components are working and if they need modification. In the case of DC Reads, information from Year 1 combined with initial data from Year 2 indicated that the use of only one of the two curricula at each school needed to be changed. Senior staff were willing to implement changes mid-year based on this information despite the challenges of re-structuring the program within a short time period. Making the changes within the current school year, rather than waiting until the following school year, allowed tutors to use the blended curriculum and to work more effectively with students.

Effective Practices

- ✓ Establishing an easy-to-use form to document student progress
- ✓ Using student SAT 9 scores as a criterion for teacher referral
- ✓ Making program modifications at mid-year

Lessons Learned

- ✓ Exploring the best mix of staff positions to monitor tutor performance and overall program quality
- ✓ Feasibility of using the SAT 9 or other group-administered tests to measure student progress
- ✓ Yopp-Singer most sensitive to gains in phonemic awareness

b. Lessons Learned

The AmeriCorps*VISTA members were responsible for the day-to-day supervision and monitoring of tutors. For example, AmeriCorp*VISTA members provided feedback to tutors regarding the quality and appropriateness of their comments in student logs. While the AmeriCorps*VISTA members have done the best job they can with the training they have received, many did not have the training and experience to work with tutors on improving instructional strategies when this was required. School reading specialists were envisioned as the individuals who could provide this kind of support. Given that reading specialists were only involved with DC Reads in some schools, the CIS reading specialist/trainer assumed responsibility for monitoring the quality of some tutoring sessions. This approach was not realistic given her other responsibilities. The overall monitoring of tutors and program quality is not yet performed to the extent program managers believe is necessary. Further exploration of the best combination of resources to perform monitoring activities, including using school-based reading specialists, AmeriCorps*VISTA members, project site directors, and the reading specialist/trainer, will continue in Year 3 of the program.

The program evaluation was designed to determine if tutored students improved their reading skills after participating in the DC Reads tutoring program. A battery of tests was administered to students participating in the program at four schools. Students' reading scores on the SAT 9, which is taken by all students each fall and spring, were also examined to see if the results were similar to those from the individually administered tests in the performance battery. The correlational analysis between the performance battery of tests and the SAT 9 found that the SAT 9 scores could be used in lieu of individual test scores for future evaluations of DC Reads or other similar tutoring programs, provided the use of these scores was limited to second grade and older students. The analysis concluded that the SAT 9 detected improvement in reading skills. It was also positively correlated with most reading performance battery tests. However, of the individually administered tests in the Reading Performance Battery, the Yopp-Singer was most sensitive to gains made by tutored students.

7. Access to Training and Technical Assistance

Effective Practices

- ✓ Ability to access the school's reading specialist
- ✓ Access to training resources

Lessons Learned

- ✓ Increasing the availability of reading specialists to the program

We can draw two effective practices and one lesson learned regarding access to training and technical assistance resources from the CIS DC Reads program. The effective practices are within-school access to a reading specialist and access to the program trainer for additional reading support. The lesson learned was that while in theory a school-based reading specialist was available to the program, few tutors had access to this individual within their school.

a. Effective Practices

In the original program design, the DC Reads program incorporated the involvement of the reading specialist at each participating school. The reading specialists were to provide feedback, offer suggestions, and monitor

program quality. They were to serve as a technical resource to the AmeriCorps*VISTA members and tutors, assisting with problem solving. In schools where the reading specialist was involved, the timely response of the reading specialist was seen as great asset to the program.

To ensure that appropriate resources were available to tutors and AmeriCorps*VISTA members, particularly in the event the school reading specialists were not readily available, CIS created the trainer/program development specialist position. In addition to providing training and developing curriculum materials, the trainer also performed some monitoring of the quality of tutoring sessions. The trainer was able to offer assistance and provide retraining as it was required.

b. Lessons Learned

Although they were eager to obtain support from reading specialists, the reading specialists were not always available to tutors or AmeriCorps*VISTA members. In some instances this was because the school did not have a reading specialist; in others, it was because the reading specialist had other responsibilities that precluded participation in the DC Reads program. Reading specialists were not involved in most of the CIS DC Reads program schools although tutors indicated they desired more support from their school's reading specialist. Therefore, to the extent possible, reading specialists should be involved to provide periodic support for the program. To ensure that resources are available to tutors when the reading specialist is unable to participate, the program has the reading specialist/trainer who can provide some ongoing support.

8. Engendering Positive, Caring Relationships Among Students, Staff, and Tutors

Tutors worked to develop positive relationships with the students they tutored. Using Federal Work-Study students as tutors was also an effective practice. The lessons learned were that tutor recruitment was difficult and that tutor turnover disrupted student-tutor relationships.

a. Effective Practices

Tutors responded to the student's questions, provided encouragement, listened attentively, and exhibited enthusiasm about the child's efforts during tutoring sessions. They interacted well with the children and interactions between tutor and child were overall very positive.

The children in the program were able to spend one-on-one time with their tutors and develop a personal relationship with a caring young adult, an opportunity they may not have otherwise enjoyed while in school. Students looked forward to the tutoring and wanted to continue with their tutor even when they were no longer in the program. Tutors' relationships with the children were viewed very positively by school staff. School staff (along with parents, AmeriCorps* VISTA members, and tutors) also believed that the children showed improvement in academics and in behaviors that contribute to academic success.

Another effective practice is using Federal Work-Study students to function as tutors. These students were committed to tutoring a certain number of hours a week as part of their Work-Study requirements. Thus, for each semester they participated in the Work-Study program, they were quite consistent in working with the children.

b. Lessons Learned

Effective Practices

- ✓ Positive interactions between tutors and students
- ✓ Use of Federal Work-Study students as tutors

Lessons Learned

- ✓ Challenges in recruiting and retaining tutors
- ✓ Impact on students because of tutor turnover

Tutor recruitment and retention was an issue during Year 2 of the DC Reads program. In particular, tutor recruitment was a considerable challenge. One reason for recruitment difficulties was that many of the CIS DC Reads schools were generally not easily accessible by public transportation. A second reason was related to how potential tutors were assigned to various tutoring programs. In Year 2 of the program, there were other organizations that provided similar tutoring services. Some of the universities that provided Federal Work-Study students as tutors only to CIS DC Reads sites in Year 1 began sending their students to other tutoring sites. The demand for tutors exceeded the available number of students who could tutor. Also, in some instances the Federal Work-Study tutors' schedules changed at mid-year, precluding them from continuing to tutor for the full school year. Finally, CIS managers had anticipated being able to recruit large numbers of community volunteer tutors. In fact, they experienced difficulties working with the communities and had limited success in recruiting volunteers from them. CIS managers found that potential community volunteers for the DC Reads program were often single parents who did not have the time to volunteer. In addition, some people in the community did not have adequate levels of literacy to tutor effectively.

The program experienced a great deal of tutor turnover, which had an impact on the students. In some instances, children lost their tutor and had to start with a new tutor. Some students went for long periods without a tutor, but were picked up again when a new tutor was recruited. For others, the lack of a replacement tutor required that the child be dropped from the program. These children found it difficult to understand why their tutor did not continue to tutor them any longer.

Strategies to improve tutor recruitment and retention are currently being developed, such as approaching the private sector and government organizations (e.g., U.S. Department of Education) to recruit volunteer tutors. For schools that are not easily accessible to public transportation, program staff plan to develop additional strategies to assist them in recruiting tutors to serve at these schools.

Chapter IV. Conclusion

Students who received tutoring made significant gains in reading, and overall the program experienced considerable success in its second year. The experiences of the partners implementing the DC Reads program offer valuable lessons to others involved in planning or implementing similar programs. In Year 3, the partners expect to continue to build on previous successes while addressing the remaining challenges.

A. Achievements Related to Student Outcomes

There is strong evidence that children with low reading performance can benefit from a program such as the one implemented by the DC Reads Collaborative.

1. Reading Performance Test Battery

Students demonstrated statistically significant gains on five of the eight measures in the battery. There were impressive gains on the measure of phonemic awareness. The standard deviation decreased on six of the eight measures, indicating that students were performing not only at a higher level but also more consistently as a group at the end of tutoring. This means that fewer students were falling far behind. In addition, on three of the eight measures, students who began the year with lower skills had a faster growth rate than those who entered with higher skills. This indicates that students were able to gain ground in some areas.

2. Stanford Achievement Test

Tutored students demonstrated greater gains on the SAT 9 reading tests than did nontutored students. While the gains for both groups were significant, tutored students gained nearly twice as much as nontutored students, thus narrowing the gap between lower performing tutored students and their higher performing nontutored classmates.

3. Language Arts Grades

The grades of all second grade students increased from first to fourth quarter in both reading and writing. The strongest gains were among those tutored students who were not tested individually for the evaluation.

4. Perceived Academic and Personal Benefits

AmeriCorps*VISTA members, tutors, teachers, principals, and parents all reported that the children who participated in DC Reads showed improvement in academics and in behaviors that contribute to academic success. Students also appeared to benefit personally from the opportunity to spend one-on-one time with a supportive young adult.

B. Implications for Other Tutoring Programs Focused on Reading Improvement

There is much to be learned from the achievements and challenges faced by the partners who implemented DC Reads for two years. Below we present guidelines for others to use in planning and refining similar programs.

1. Involvement of the School District and Participating Schools

Support from the central school district office is very important (particularly in the early planning and implementation stages), but the support of principals is key to a successful implementation.

Teachers should be involved in the process of selecting students to participate. This will ensure that the students most in need receive tutoring and it will increase teacher support for the program.

School staff, including the teachers of students who receive tutoring, should be informed early on about the program in general and about the curriculum materials and activities used in tutoring sessions.

Programs should encourage the involvement of school staff, but build in back-up systems for training and technical assistance. In particular, because reading specialists have other full-time responsibilities, it is generally not advisable to count on their active involvement.

2. Scheduling and Facilities

The schedule for tutoring should accommodate the needs of the school. In particular, it should not interfere with regular classroom instruction in core subjects.

Recognize that the actual session length will be somewhat less than the scheduled time. Allow 10 to 15 minutes for logistics and communication.

The facilities available to tutors and their supervisors (in this case, AmeriCorps*VISTA members) should be comfortable and adequate in size for the maximum number of child/tutor pairs at the site.

3. Tutor Recruitment and Retention

Recruiting tutors is a major challenge and should not be underestimated. The result of significant problems in tutor recruitment is that children receive many fewer than the expected number of sessions.

Retaining tutors is also a problem, particularly with Federal Word Study students who may experience dramatic schedule changes from one semester to the next. When tutors drop out,

children lose both the opportunity to progress academically and the warm, supportive relationship that has been established between tutor and child.

4. Tutor Training, Supervision, and Technical Assistance

Training for tutors must be readily available, in part due to the need to provide continuous training support for newly recruited tutors. Programs should develop their own internal training capability rather than relying on school system staff.

AmeriCorps*VISTA members can provide appropriate day-to-day logistical supervision to tutors and are likely to be well regarded by professional school staff and by tutors.

Technical assistance must be available to tutors and their supervisors. AmeriCorps*VISTA members have only limited training and education in reading and may not be able to address technical problems that arise. Technical assistance could be delivered by school-based reading specialists if they are available and willing, by the programs' training specialist, or by project site directors if they possess the necessary expertise in reading.

5. Instructional Materials

A range of high-quality, tested curriculum materials should be available for use in tutoring sessions. Having materials appropriate for various skill levels will allow tutors, who have received minimal training, to address the needs of all children referred to the program.

Materials used by tutors (e.g., lesson plans, progress logs) should be highly structured and easy to use. In particular, highly structured lesson plans will allow tutors with various strengths and experiences to provide appropriate and effective instruction to tutored students.

6. Evaluation

The tutoring program should include systematic progress monitoring and assessment. This should include both a record of each student's progress that documents the difficulties the child encountered and regularly scheduled formal assessment.

Ongoing evaluation of the program is critical. The evaluation should include assessments of both implementation processes and student outcomes. In the case of student outcomes, group tests administered by the school districts such as the SAT 9 offer a low-cost available source of information. Based on the Year 2 evaluation of DC Reads, it appears that such tests are appropriate, at least for students who are in the second or higher grades. However, they are not as sensitive to gains in phonemic awareness as the specialized, individually administered Yopp-Singer Test of Phonemic Segmentation.

7. Communication

Mechanisms should be in place to encourage ongoing communication between the tutors and AmeriCorps*VISTA members and the classroom teachers about progress in tutoring sessions and in the classroom. While coordination of tutoring and classroom instruction is probably unrealistic, communication will help to ensure that tutors and teachers are not working at cross purposes and that they share a common understanding about children's needs.

When organizations with disparate cultures come together, the communication challenges must be recognized and dealt with openly. Misunderstandings are particularly likely to occur when the partnership involves organizations with distinctly different racial/ethnic or corporate cultures.

C. Next Steps for DC Reads

In 2 years, the DC Reads partners have come a long way together. Overall, the CIS DC Reads program appeared to be operating well. CIS made improvements to the program during Year 2. Of particular note were improvements in the tutoring curriculum and in the training provided to AmeriCorps*VISTA members and tutors. In general, school staff were quite satisfied with the DC Reads program and wanted it to continue to operate in their schools. Most of those working and volunteering in the program were also satisfied with the program and their experiences in being associated with it.

Together, the individual and group testing provided evidence that tutoring by volunteers with minimal training was able to improve the reading skills of low-performing children. This finding is particularly noteworthy because the intensity of the tutoring was not as high as had been hoped. Tutored, tested students received, on average, just 20 sessions—a number that according to previous research by others, is at the low end of what is needed to obtain improvement. It seems likely that with more sessions, the students in this study would have demonstrated even greater gains. It is also worth noting here that the schools selected for the outcome study appear to have experienced greater difficulties recruiting and retaining tutors than most other CIS sites. It seems likely that in some other schools, students may have received more sessions and improved even more than our data indicate. We consider the outcome study results to be a conservative estimate of the potential impact of this program on students' reading performance.

This program, with its limited number of sessions, was most successful at improving phonemic awareness and basic decoding. This suggests that the curricula used are effective in developing the basic skills that are prerequisite to comprehension. It was less successful at improving word recognition and reading comprehension. We hypothesize that increasing the number of sessions would produce better results in word recognition and comprehension, enabling students to move beyond basic decoding and into proficient reading.

Challenges remain, and they will have to be dealt with as the DC Reads model evolves. The most noteworthy are challenges related to 1) tutor recruitment and retention, and 2) communication among partners. Despite the challenges, many partners were pleased with the

program's achievements and the partnership that had been formed. As one partner representative commented, "One of the things that is most exciting is that the partners have been willing to revisit all assumptions and decisions to make the program work. They are flexible and willing to continuously evaluate and improve the program. They are willing to explore new models and experiment—and we are coming up with stronger model."

Appendix A: Methodology

A. Evaluation Components

The Year 2 evaluation of DC Reads conducted by Macro International included assessments of program implementation, achievements, and outcomes.

1. Implementation Assessment

The purpose of the Implementation Assessment was to provide 1) information for program improvement, 2) a context for interpreting the results of the achievement and outcome data, and 3) lessons learned for other programs that provide tutoring in reading.

2. Investigation of Program Achievements

Macro examined the achievements of the DC Reads program in terms of number of students tutored, tutors recruited, and hours of tutoring provided. We also considered the opinions of people associated with the program regarding the achievements of DC Reads and their satisfaction with the program.

3. Outcome Assessment

The outcome assessment was a study that focused on tutored students' reading performance. The assessment included examination of the relationships between students' performance on an individually administered Reading Performance Battery of tests with 1) their scores on Stanford Achievement Test reading subtests and 2) their language arts grades. The outcome assessment included an investigation of variations in reading skill growth and aspects of the tutoring program that contributed to reading improvement.

B. Study Schools and Students

This Year 2 evaluation of DC Reads focused on programs operated by Communities in Schools (CIS). CIS operates DC Reads programs in 16 schools designated by DC Public Schools in 1997 as the schools in most need of academic assistance. Four of those schools were selected as primary study sites. Within the four schools, 85 second-grade tutored students were randomly selected from all tutored students as the primary subjects for the outcome study.

Based on our analysis of fall 1998 Stanford Achievement Test (SAT 9) scores of second graders in the four study sites, students selected for tutoring were performing at significantly lower levels in reading than their nontutored peers. There was no significant difference in reading performance between tutored students who participated in the outcome study and those who did not.

C. Assessment Schedule for Students Participating in the Outcome Study

The evaluation design called for students participating in the outcome study (i.e., tutored, tested students) to be tested through administration of the Reading Performance Battery on three occasions during the school year. The overall design is as follows:

Wave 1 testing — Begin tutoring — Wave 2 testing — End of tutoring — Wave 3 testing

Wave 1 testing took place in October and November 1998 (N=85). Wave 2 testing took place in January and February 1999 (N=75). Wave 3 testing took place in May 1999 (N=51).

D. Data Sources

Sources of data for the evaluation included:

A Reading Performance Battery of individually administered tests

DC Public Schools data (including, SAT 9 reading scores and language arts grades)

Data collection instruments developed for the Year 2 evaluation (i.e. interview and observation protocols, survey questionnaires, and recordkeeping logs).

1. Reading Performance Assessment

Six education test specialists, each rotating among the four study schools, administered a one-on-one battery of reading tests to tutored students during the time set aside for them to attend tutoring sessions. Most tests on the Reading Performance Battery were administered three times during the school year. Tests in the Battery were:

1) Identification of Capital Letters and Lower Case Letters
2) Three tests from the Woodcock Reading Mastery Tests—Word Identification, Word Attack, and Passage Comprehension
3) Yopp-Singer Test of Phoneme Segmentation
4) Two components of the Oral Reading subtest of the Durrell Analysis of Reading Difficulty (Oral Reading Time and Oral Reading Comprehension).

The content and administration of the tests in the Reading Performance Battery is more fully described in Chapter I of this report.

2. Stanford Achievement Test Scores

The Stanford Achievement Test (SAT 9) is routinely administered to all DCPS second-grade students in the fall and spring of each school year. Macro obtained fall and spring SAT 9 scores

for all second graders in the four study schools in the following areas: Word Study Skills, Word Reading Skills, Reading Comprehension, and Total Reading.

3. Language Arts Grades

Macro researchers obtained reading and writing grades of tutored and nontutored second-grade students for the first and fourth quarters from all four study schools.

4. Observations

At two points during the school year, in November 1998 and March 1999, Macro researchers visited each of the four study schools and conducted observations of tutoring sessions. In all, we observed 21 tutor-child pairs for a full tutoring session. Sessions were scheduled to last for either 40 minutes or 60 minutes. The protocol used to conduct the observations is included in Appendix C, Instruments.

The observation protocol is composed of seven sections.

The first captured basic identifying information about the school and tutor, as well as session start and end times.

The second section captured some information about the student, including number of sessions to date, starting level, and current level. (This information was drawn from the Student Progress Log by the observer.)

The next section included the amount of time spent on specific activities during the various phases of the tutoring session. Phases for a student using Book Partners materials were rereading familiar material, reading an Early Success book, and reading a Choice Time book. Phases for a student using Reading One-to-One materials were review, direct instruction, reading, and writing.

The fourth section was used to record the amount of time that various materials (e.g., books, letter tiles) are used.

The fifth section captured information about tutor-child interactions, focusing primarily on tutor behaviors.

The sixth section was used to assess the appropriateness of the instructional level.

The final section provided space for additional comments (e.g., about the tutoring environment, the child's overall behavior, session logistics).

5. Interviews

During the same two time periods that Macro researchers observed tutoring sessions in schools, the researchers also interviewed principals, second-grade teachers, and AmeriCorps*VISTA members. Typically, in schools with more than one AmeriCorps*VISTA member, the two were interviewed together. The principal and teacher interviews focused on satisfaction with:

1) Performance of the AmeriCorps*VISTA member(s)
2) Performance of the tutors
3) Implementation and administration of DC Reads in the school.

The AmeriCorps*VISTA member interviews focused on satisfaction with

1) Support received from the school
2) Support received from the program
3) Performance of tutors
4) Implementation and administration of DC Reads in the school.

6. Surveys

Macro developed survey questionnaires for the following groups of individuals involved in DC Reads:

AmeriCorps*VISTA members
Tutors
Principals
Teachers
Parents
Project site directors
Campus coordinators.

AmeriCorps*VISTA member, tutor, principal, teacher, and parent surveys were piloted in winter 1999. The questionnaires that were distributed in study schools through AmeriCorps*VISTA members or campus coordinators (for some tutors). Respondents were directed to return surveys directly to Macro in a postage-paid self-addressed envelope. Macro conducted an initial mailing, sent a reminder notice, and completed two full remailings to all nonrespondents.

Response rates for the survey questionnaires were as follows: AmeriCorps*VISTA members—7 of 7, or 100 percent; tutors—14 of 30, or 47 percent; principals—4 of 4, or 100 percent; teachers—5 of 13, or 38 percent; parents—9 of 52, or 17 percent; campus coordinators—3 of 3, or 100 percent; project site directors—3 of 3, or 100 percent.

The questionnaires used in the survey are included in Appendix C, Instruments. Areas of interest on the questionnaires include:

1) Position responsibilities (asked of AmeriCorps*VISTA members, tutors, project site directors, and campus coordinators)
2) Training received (asked of AmeriCorps*VISTA members and tutors)
3) Performance of AmeriCorps*VISTA members (asked of tutors, principals, teachers, and project site directors)
4) Performance of tutors (asked of AmeriCorps*VISTA members, principals, teachers, and parents)

5) Communication and relationships within DC Reads schools (asked of AmeriCorps*VISTA members, tutors, principals, and teachers)

6) Communication and relationships among partners (asked of project site directors and campus coordinators)

7) Satisfaction with the overall program model and materials (asked of all respondents)

8) Improvement in students' performance and behavior (asked of AmeriCorps*VISTA members, tutors, principals, teachers, and parents)

9) Areas in which DC Reads has been successful or needs improvement (asked of all respondents).

7. Student Progress Logs

A progress log was maintained for each child. Similar, though slightly different, logs existed for the two curricula (Book Partners and Reading One-to-One), and for the different levels within the Reading One-to-One curriculum (Alphabet, Word Family, and Reading Comprehension). During the second semester, once the two curricula were merged, the Book Partners log was adopted at all four study schools. Macro researchers monitored logs throughout the year to ensure that they were complete and legible. At the end of the school year, Macro obtained copies of all logs for analysis. Elements of the log of interest for the study were: total length of session in minutes, book(s) used, activities of the session (e.g., review, reading, working with words, and writing), and attentiveness of the student (i.e., poor, average, or excellent).

8. Logs of Tutors and Tutored Students

Macro researchers attempted to gather tutor and student logs from all 16 CIS sites. These logs document start and end dates and reasons for joining or leaving the program. Macro was able to obtain tutor logs from 10 schools, including 3 study schools, and student logs from 15 schools, including all study schools.

Appendix B: Sources of Information

Table B-1 presents the sources of information on which we have based the effective practices and lessons learned presented in Chapter 3 of this report.

Table B-1: Effective Practices and Lessons Learned By Key Program Components

Item	Survey	Interview	Tutoring Observation	Logs	Other Sources
Key Program Component 1: Use of research-based elements to produce reading achievement					
Use of well-tested curricula developed at major universities (EP[1])		Corporation Staff CIS Staff			Reading Performance Battery Stanford Achievement Test 9 Curriculum Documentation
Blending two curricula to better meet student needs (LL[2])		CIS Staff Corporation Staff AmeriCorps*VISTAs		Progress	
May still require additional materials to meet the needs of some children (LL)	Tutors Principals Teachers	CIS Staff Corporation Staff			
Key Program Component 2: Well-structured tutoring session in which the content and delivery of instruction are carefully planned					
A curriculum that includes a wide range of materials to meet the varied needs of students (EP)	Tutors	CIS Staff Corporation Staff	Fall Spring		
A corresponding lesson plan for each book (EP)	Tutors	CIS Staff Corporation Staff	Fall Spring		Lesson Plans
Developing individual lesson plans rather than using large curriculum manuals (LL)	Tutors	CIS Staff Corporation Staff	Spring		Reading One-to-One Manual
Recognizing the need to develop highly structured lessons to guide tutors (LL)		CIS Staff	Fall Spring		Comparison of Year 2 observations with Year 1 observations

[1] EP indicates that this is an effective practice.
[2] LL indicates that this is a lesson learned.

BEST COPY AVAILABLE

Item	Survey	Interview	Tutoring Observation	Logs	Other Sources
Key Program Component 3: Close coordination with schools, school administration, and the classroom and/or reading teacher					
Involving teachers in selecting students for tutoring (EP)		CIS Staff AmeriCorps*VISTAs Principals Teachers			Teacher Form: Selection of Students
Involving the central public schools (superintendent's) office (EP)		CIS Staff Corporation Staff			
Actively involving school principals (EP)	Project Site Directors AmeriCorps*VISTAs Principals	CIS Staff AmeriCorps*VISTAs Principals			
Having an on-site AmeriCorps*VISTA member as a coordinator at each school (EP)	Project Site Directors Principals Teachers	CIS Staff Corporation Staff Principals Teachers			
Providing information about the program at the beginning of the school year (EP)	Project Site Directors AmeriCorps*VISTAs Principals Teachers	AmeriCorps*VISTAs Principals Teachers			
Improving access to facilities and equipment (LL)	Project Site Directors AmeriCorps*VISTAs Tutors	AmeriCorps*VISTAs			
Increasing communication with teachers (LL)	Project Site Directors AmeriCorps*VISTAs Tutors Teachers	AmeriCorps*VISTAs Teachers			
Increasing access to reading specialists (LL)	AmeriCorps*VISTAs Tutors	AmeriCorps*VISTAs			
Recognizing how cultural differences impact communication (LL)		CIS Staff Corporation Staff			
Key Program Component 4: Intensive and ongoing training and supervision for tutors					
Establishing a training coordinator position (EP)	AmeriCorps*VISTAs Tutors	CIS Staff Corporation Staff			
Designating AmeriCorps*VISTA members as tutor supervisors (EP)	Project Site Directors Tutors	CIS Staff Corporation Staff	Fall Spring		

BEST COPY AVAILABLE

Item	Survey	Interview	Tutoring Observation	Logs	Other Sources
	Principals Teachers				
Not relying on existing public school staff to provide timely training (LL)		CIS Staff Corporation Staff			
Key Program Component 5: Frequent and regular tutoring sessions					
Establishing specific days and times for tutoring sessions (EP)	AmeriCorps*VISTAs Tutors Principals Teachers	CIS Staff Corporation Staff AmeriCorps*VISTAs Principals Teachers	Fall Spring		
Increasing the number of sessions in which tutored children participate (LL)			Fall Spring	Progress Student Tutor	
Ensuring tutoring sessions are held for the designated length of time (LL)			Fall Spring		
Reconciling tutoring session length between the two curricula (LL)		CIS Staff Corporation Staff			

BEST COPY AVAILABLE

Item	Survey	Interview	Tutoring Observation	Logs	Other Sources
Key Program Component 6: Careful evaluation, assessment, monitoring, and reinforcement of progress					
Establishing an easy-to-use form to document student progress (EP)		CIS Staff Corporation Staff AmeriCorps*VISTAs		Progress	Reading One-to-One Progress Logs
Using student SAT 9 scores as a criterion for teacher referral (EP)		CIS Staff Corporation Staff AmeriCorps*VISTAs			Teacher Form: Selection of Students Stanford Achievement Test 9
Making program modifications at mid-year (EP)		CIS Staff Corporation Staff AmeriCorps*VISTAs			
Exploring the best mix of staff positions to monitor tutor performance and overall program quality (LL)	AmeriCorps*VISTAs Tutors	CIS Staff Corporation Staff			
Feasibility of using the SAT 9 or other group-administered tests to measure student progress (LL)					Stanford Achievement Test 9 Reading Performance Assessment School Documentation
Key Program Component 7: Access to training and technical assistance resources					
Ability to access the school's reading specialist (EP)	AmeriCorps*VISTAs Tutors	CIS Staff Corporation Staff AmeriCorps*VISTAs			
Access to training resources (EP)	AmeriCorps*VISTAs Tutors	CIS Staff Corporation Staff AmeriCorps*VISTAs			
Increasing the availability of reading specialists to the program (LL)	AmeriCorps*VISTAs Tutors	CIS Staff Corporation Staff AmeriCorps*VISTAs			

Item	Survey	Interview	Tutoring Observation	Logs	Other Sources
Key Program Component 8: Engendering positive, caring relationships among students, staff, and tutors					
Positive interactions between tutors and students (I,P)	Project Site Directors AmeriCorps*VISTAs Tutors Principals Teachers Parents	AmeriCorps*VISTAs Principals Teachers	Fall Spring		
Use of Federal Work-Study students as tutors (I,P)	Project Site Directors AmeriCorps*VISTAs Tutors Principals Teachers Parents	CIS Staff Corporation Staff AmeriCorps*VISTAs Principals	Fall Spring		
Challenges in recruiting and retaining tutors (I,L)	Campus Coordinator Project Site Director AmeriCorps*VISTAs Principals	CIS Staff Corporation Staff AmeriCorps*VISTAs		Progress Student Tutor	
Impact on students because of tutor turnover (I,L)		AmeriCorps*VISTAs			

Appendix D: Outcomes

The purpose of the outcome assessment was to determine if tutored students improved their reading skills and attitude toward reading after participating in the DC Reads tutoring program. The battery of tests administered for this study is described in the main body of this report.

To summarize, Macro testers assessed the reading skills of randomly selected tutored students in the four study schools. Most tests were administered three times. The exceptions were two subtests of the Durrell Analysis of Reading Difficulty. This test was administered only twice, once prior to the commencement of tutoring and once after tutoring had ended for the year. In addition, AmeriCorps*VISTA members administered the Primary Reading Index, a measure of attitudes toward reading, at the beginning and end of the school year.

We also examined students' Stanford Achievement Test (SAT 9) scores to see if the results were similar to those from the individually administered tests in the performance battery. In particular, we investigated the correlation between Stanford 9 subtests and the various tests in the battery.

Finally, we obtained student grades in reading and writing and school attendance records for all second graders. We examined grades to determine if tutored students' reading grades improved from the first to fourth quarter, and if tutored students fourth quarter grades were comparable to those of nontutored students. We examined attendance data to determine how the attendance of tutored students compared with that of nontutored peers.

A. Changes in Student Reading Performance: Comparison of Wave 1 and Wave 3

Macro examined the Reading Performance Battery results for Wave 1 (administered in October and November 1998) and Wave 3(administered in May 1999) to ascertain whether there were significant gains in students' reading performance. The following section summarizes the results. We begin by presenting Table D-1, which compares the mean scores and score changes for six of the eight tests in the battery (tests whose scores are a continuous variable, either a percentage or Standard Score[1]). Table D-1 also shows the number and percentage of students who showed improvement on Wave 3 tests.

---

[1] The Woodcock Reading Mastery Test yields many types of scores for its subtests, among them the Word Identification, Word Attack, and Passage Comprehension subtests used in this study. Standard scores have a mean of 100 and a standard deviation of 15.

Table D-1: Change in Tutored, Tested Students on Reading Performance Battery From Wave 1 to Wave 3[2]										
Reading Performance Battery Test	Testing	Total Mean	N	Std	Changes Mean	Changes Std	t	p	N Improved	% Improved
Percentage of Capital Letters	Wave 1	**98.72**	51	2.95						
	Wave 3	**99.77**	51	0.91	**1.06**	2.89	**2.61**	**0.01**	11	22%
Percentage of Lower Case Letters	Wave 1	95.93	51	4.03						
	Wave 3	96.23	51	5.57	0.30	5.32	0.40	0.69	18	35%
Percentage of Yopp-Singer	Wave 1	**52.67**	51	28.40						
	Wave 3	**76.56**	51	15.00	**23.89**	28.34	**6.02**	**0.00**	40	78%
Word Identification Standard Score	Wave 1	88.82	51	10.76						
	Wave 3	88.41	51	7.93	-0.41	5.54	-0.53	0.60	18	35%
Word Attack Standard Score	Wave 1	**89.75**	51	12.97						
	Wave 3	**92.29**	51	11.75	**2.55**	8.99	**2.03**	**0.05**	32	63%
Passage Comprehension Standard Score	Wave 1	86.65	51	12.39						
	Wave 3	88.33	51	9.80	1.69	10.09	1.19	0.24	26	51%

(Statistically significant differences are in boldface.)

Students showed statistically significant change on three tests:

Capital Letters. In both Wave 1 and Wave 3 testing, the tutored, tested students demonstrated knowledge of nearly all Capital Letters. Despite the ceiling effect demonstrated in Wave 1, 11 out of 51 students improved on Wave 3 testing, and the spread of students' scores narrowed greatly. (With p=.01, there is a 99 percent chance that this significant difference was not due to chance.)

Yopp-Singer Test of Phoneme Segmentation. The tutored, tested students improved greatly in phonemic awareness, with the mean score increasing from 52.67 percent to 76.56 percent. Most students (40 out of 51) improved their phonemic awareness from Wave 1 to Wave 3 testing. (With p=.00, there is virtually no chance that this significant difference was due to chance.)

---

[2] Many of the tables presented in this appendix present mean, N, standard deviation, t statistics, and p values.

The mean is simply the average of scores or other numerical data (e.g., grades, days absent). It is obtained by adding up the values and dividing by the number of cases.

N is simply the number of subjects for whom we have data on a particular variable (e.g., test score).

The standard deviation is a statistic that shows the spread or dispersion of scores. A high standard deviation indicates that there is a great deal of variation in scores, while a low standard deviation indicates that there is little variation.

t is the statistic generated in t tests. A t test is often used, as here, in comparing the average scores for one group with the average scores for another. The higher the magnitude of the t value, the more likely it is that a statistically significant difference between the groups exists.

The p value gives the probability that a difference between the two group means might have been due to chance. In most studies, if p=.05, the difference is considered to be a real difference and not attributed to chance.

BEST COPY AVAILABLE

Word Attack. Students' scores also increased from Wave 1 to Wave 3 on the Word Attack subtest of the Woodcock Reading Mastery Test. Mean scores increased from 89.75 in Wave 1 to 92.29 in Wave 3. (With p=.05, there is a 95 percent chance that this significant difference was not due to chance.)

In addition to tests whose scores can be expressed as a percentage or standard score, students also improved on the two subtests of the Durrell Analysis of Reading Difficulty—Oral Reading Time and Oral Reading Comprehension. The way in which these tests are administered and scored is described in main body of this report. Table D-2 presents the results.

From Wave 1 to Wave 3, tutored, tested students improved significantly from a mean of 3.10 in Oral Reading Time (Z=5.08, p=.00). This means that students were able to proceed further through the test, meeting the test's requirements for reading rate and accuracy. In terms of grade equivalents, students' median grade equivalent moved from first-grade middle (1M) to second grade low (2L).

Comprehension also improved significantly from a mean of 2.71 in Wave 1 to a mean of 3.00 in Wave 3. In addition, while students ranged from Poor to Good in comprehension in Wave 1 testing, all students had Good comprehension of what they read in Wave 3 testing.

Table D-2: Durrell Analysis of Reading Difficulty

Subtest		Mean	N	Std	Z	p	Median	Range
Durrell Reading Time	Wave 1	**3.10**	51	1.60			1M	1.6-2M
	Wave 3	**4.76**	51	2.16	**5.08**	**0.00**	2L	1.6-3M
Durrell Comprehension	Wave 1	**2.71**	51	0.61			Good	Poor-Good
	Wave 3	**3.00**	50	0.00	**2.92**	**0.00**	Good	Good-Good

(Statistically significant differences are in boldface.)

The results for all tests in the Reading Performance Battery can be summarized as follows:

In Wave 3 testing the tutored, tested students showed significant improvement in identification of Capital Letters, Phonemic Segmentation, and Word Attack. They also performed better on measures of Oral Reading Time and Oral Reading Comprehension. There were no significant declines in scores.

The percentage of students who show improvement across the six tests on the Reading Performance Battery ranges from 22 percent to 78 percent. More than half of the students showed improvement on tests of Phonemic Segmentation, Word Attack, and Passage Comprehension.

With the exception of Lower Case Letters and Oral Reading Time, the standard deviations on all tests decreased from Wave 1 to Wave 3. For Capital Letters, Phonemic Segmentation, Word Identification, Word Attack, Passage Comprehension, and Oral Reading

Comprehension the spread in tutored, tested students' scores decreased. This indicates that tutored students performed more consistently as a group after being tutored.

B. Stanford Achievement Test Results

Tutored students made greater gains on the Stanford Achievement Test (SAT 9) reading tests than nontutored students did. On Normal Curve Equivalent (NCE) scores—which compare students to all other students in the nation at the beginning and end of the school year—tutored students gained nearly twice as much as nontutored students. Also, 69 percent of tutored students improved on Total Reading, while just 60 percent of their nontutored peers demonstrated improvement.

Mean NCE[3] scores of tutored students on Total Reading increased significantly and substantially more than scores of nontutored students. Tutored students gained nearly 7 points, and a gain of 5 points is considered significant by the test's publisher.

Tutored, tested students went from 28.17 to 33.10, a gain of 4.93
Tutored, nontested students went from 29.04 to 38.55, a gain of 9.50.
Nontutored students went from 45.31 to 48.87, a gain of 3.56

Table D-3 details the results on the SAT 9 for tutored and nontutored students.

Table D-3: Stanford Achievement Test Normal Curve Equivalent Scores

SAT 9 Test	Test Date	Mean	N	Std	Changes Mean	Changes Std	t	p	N Improved	% Improved
Tutored, Tested Students										
Total Reading	Fall	**28.17**	42	9.19	**4.93**	8.20	**3.89**	**0.00**	29	69%
	Spring	**33.10**	42	12.19						
Word Study Skills	Fall	33.30	46	9.11	1.88	11.10	1.15	0.26	27	59%
	Spring	35.18	46	11.41						
Word Reading	Fall	28.86	45	12.62	1.00	13.61	0.49	0.63	27	60%
	Spring	29.86	45	15.63						
Reading Comprehension	Fall	**28.74**	43	10.34	**6.51**	9.87	**4.33**	**0.00**	30	70%
	Spring	**35.26**	43	13.11						

[3] Normal Curve Equivalent (NCE) scores have a mean of 50 and a standard deviation of 21.06.

Table D-3 (continued): Stanford Achievement Test Normal Curve Equivalent Scores										
SAT 9 Test	Test Date	Mean	N	Std	Changes Mean	Changes Std	t	p	N Improved	% Improved
Tutored, Nontested Students										
Total Reading	Fall	**29.04**	29	10.40	**9.50**	13.78	**3.71**	**0.00**	20	69%
	Spring	**38.55**	29	15.74						
Word Study Skills	Fall	**32.53**	30	12.07	**7.87**	14.58	**2.96**	**0.01**	22	73%
	Spring	**40.40**	30	15.30						
Word Reading	Fall	30.18	29	13.44	4.53	18.02	1.36	0.19	15	52%
	Spring	34.71	29	18.40						
Reading Comprehension	Fall	**29.91**	30	10.58	**10.04**	16.37	**3.36**	**0.00**	19	63%
	Spring	**39.95**	30	18.57						
Nontutored Students										
Total Reading	Fall	**45.31**	86	16.25	**3.56**	9.45	**3.49**	**0.00**	52	60%
	Spring	**48.87**	86	18.67						
Word Study Skills	Fall	**46.28**	91	16.66	**4.26**	12.71	**3.19**	**0.00**	55	60%
	Spring	**50.54**	91	17.90						
Word Reading	Fall	46.49	87	17.32	0.89	16.31	0.51	0.61	49	56%
	Spring	47.38	87	23.39						
Reading Comprehension	Fall	**46.22**	91	16.79	**3.30**	12.82	**2.46**	**0.02**	49	54%
	Spring	**49.53**	91	17.53						
All Students										
Total Reading	Fall	**37.72**	157	15.99	**5.02**	10.26	**6.13**	**0.00**	101	64%
	Spring	**42.74**	157	17.96						
Word Study Skills	Fall	**40.24**	167	15.57	**4.25**	12.73	**4.32**	**0.00**	104	62%
	Spring	**44.49**	167	17.23						
Word Reading	Fall	38.62	161	17.61	1.58	15.90	1.26	0.21	91	57%
	Spring	40.20	161	22.00						
Reading Comprehension	Fall	**38.66**	164	16.59	**5.38**	13.05	**5.28**	**0.00**	98	60%
	Spring	**44.03**	164	17.77						

(Statistically significant differences are in boldface.)

C. Correlation of the Reading Performance Battery with the SAT 9 Reading Subtests

In this section we present our analysis of correlations between the SAT 9 reading scores of tutored, tested students and their Reading Performance Battery scores. Tables D-4 and D-5 present the correlations for Wave 1 and Tables D-6 and D-7 present the correlations for Wave 3. Continuous data and ordinal data correlations are presented separately.

Table D-4: Continuous Data Correlations—Fall '98 Stanford Achievement Tests and Wave 1 Individual Tests					
Macro Reading Performance Battery Tests		Stanford Achievement Test Reading Subtests			
		Word Study Skill	Word Reading	Reading Comprehension	Total Reading
Percentage of Yopp-Singer Items	Pearson Correlation	0.11	0.09	0.07	0.13
	Sig. (2-tailed)	0.35	0.42	0.55	0.28
	N	74	75	71	70
Word Identification	Pearson Correlation**	**0.52**	**0.65**	**0.62**	**0.71**
	Sig. (2-tailed)	**0.00**	**0.00**	**0.00**	**0.00**
	N	74	75	71	70
Word Attack	Pearson Correlation**	**0.40**	**0.54**	**0.51**	**0.58**
	Sig. (2-tailed)	**0.00**	**0.00**	**0.00**	**0.00**
	N	74	75	71	70
Passage Comprehension	Pearson Correlation**	**0.41**	**0.58**	**0.47**	**0.56**
	Sig. (2-tailed)	**0.00**	**0.00**	**0.00**	**0.00**
	N	74	75	71	70

** Correlation is significant at the 0.01 level (2-tailed).

(Statistically significant correlations are in boldface.)

Table D-5: Ordinal Data Correlations—Fall '98 Stanford Achievement Tests and Wave 1 Individual Tests					
Macro Reading Performance Battery Tests		Stanford Achievement Test Reading Subtests			
		Word Study Skill	Word Reading	Reading Comprehension	Total Reading
Durrell Reading Comprehension	Spearman Correlation	0.22	0.28*	0.39**	0.33**
	Sig. (2-tailed)	0.06	**0.02**	**0.00**	**0.01**
	N	73	**74**	**70**	**69**
Durrell Reading Time	Spearman Correlation	0.22	**0.39****	**0.33****	**0.40****
	Sig. (2-tailed)	0.06	**0.00**	**0.01**	**0.00**
	N	73	74	70	69

* Correlation is significant at the .05 level (2-tailed).

** Correlation is significant at the .01 level (2-tailed).

(Statistically significant correlations are in boldface.)

Table D-6: Continuous Data Correlations—Spring '99 Stanford Achievement Tests and Wave 3 Individual Tests					
Macro Reading Performance Battery Tests		Stanford Achievement Test Reading Subtests			
		Word Study Skill	Word Reading	Reading Comprehension	Total Reading
Percentage of Yopp-Singer Items	Pearson Correlation	0.27	0.08	0.08	0.18
	Sig. (2-tailed)	0.08	0.60	0.61	0.26
	N	43	43	43	41
Word Identification	Pearson Correlation**	**0.56**	**0.70**	**0.56**	**0.72**
	Sig. (2-tailed)	**0.00**	**0.00**	**0.00**	**0.00**
	N	43	43	43	41
Word Attack	Pearson Correlation**	**0.57**	**0.45**	**0.41**	**0.56**
	Sig. (2-tailed)	**0.00**	**0.00**	**0.01**	**0.00**
	N	43	43	43	41
Passage Comprehension	Pearson Correlation**	**0.58**	**0.77**	**0.54**	**0.73**
	Sig. (2-tailed)	**0.00**	**0.00**	**0.00**	**0.00**
	N	43	43	43	41

** Correlation is significant at the 0.01 level(2-tailed).

(Statistically significant correlations are in boldface.)

Table D-7: Ordinal Data Correlations—Spring '99 Stanford Achievement Tests and Wave 3 Individual Tests					
Macro Reading Performance Battery Tests		Stanford Achievement Test Reading Subtests			
		Word Study Skill	Word Reading	Reading Comprehension	Total Reading
Durrell Reading Comprehension	Spearman Correlation[1]	.	.	.	.
	Sig. (2-tailed)	.	.	.	.
	N	42	42	42	40.00
Durrell Reading Time	Spearman Correlation	**0.32***	**0.40****	**0.49****	**0.50****
	Sig. (2-tailed)	**0.04**	**0.01**	**0.00**	**0.00**
	N	43	43	43	41.00

1 No correlation is possible because all students scored "good" on Durrell Comprehension.

* Correlation is significant at the .05 level (2-tailed).

** Correlation is significant at the .01 level (2-tailed).

(Statistically significant correlations are in boldface.)

In both Waves 1 and Wave 3 testing, there are moderate and statistically significant correlations between most Reading Performance Battery tests and most SAT 9 subtests. The most notable exception is the Yopp-Singer Test of Phoneme Segmentation, which does not correlate strongly with any of the SAT 9 subtests. (The Yopp-Singer test is the measure on which students demonstrated most dramatic gains.) The Yopp-Singer appears to be a very sensitive measure, but does not measure a wide variety of reading skills. While phonemic awareness is a key underpinning for successful reading, it is a beginning step. The SAT 9 Word Reading and Word Study tests are probably more similar in terms of the skills they assess to the Woodcock Word Identification and Word Attack tests than they are to the Yopp-Singer.

Although there is no one-for-one match between tests in the Reading Performance Battery and SAT 9 subtests, both seem capable of detecting improvement in reading skills among second graders. The consistent correlations seem to indicate that Stanford 9 test scores could be used in lieu of individual test scores in future evaluations of DC Reads or other similar tutoring programs. A caveat is that this study dealt with second graders. For evaluations focusing on younger children, the SAT 9 scores (or any other group administered assessment) may not be suitable.

D. Attitudes Toward Reading

The Primary Reading Index, the attitude measure used for this study, consists of 34 items. The items are statements about books or reading, to which the child responds by circling a sad, neutral, or smiling face. Points are assigned on the following basis: 1 for a "sad face," 2 for a "neutral face," and 3 for a "happy face."

We examined the reliability and validity of the Reading Attitude Assessment analyzed the data to determine whether tutored students changed their attitudes toward reading during their participation in DC Reads tutoring.

To summarize, the results of our analyses of the reading attitude assessment indicate the following.

The average reading attitude score for all students is approximately 2.5 on a 3-point scale, that is midway between a neutral and smiling face. This indicates that tutored, tested students began with quite positive attitudes toward reading.

There is no evidence in the current evaluation to support the concurrent, predictive, or known group validity of the instrument used.

There is no evidence to support any reading attitude change from the beginning to the end of participating in DC Reads.

Because of the above results, the reading attitude measure was not investigated further in the Hierarchical Linear Model (HLM) growth model presented later in this appendix.

E. Language Arts Grades

The average grades in both reading and writing of all students increased somewhat from the first to fourth quarters. There was significant gain for nontutored students and for tutored, nontested students. Nontutored students had significantly higher scores than tutored students in both fall and spring testing. Table D-8 presents these results.

Table D-8: First and Fourth Quarter Language Arts Grades

Academic Area	Quarter	Mean	N	Std	Mean Change	Std Change	t	p	N Improved	% Improved
Tutored, Tested Students										
Reading Grades	First	1.80	45	0.69						
	Fourth	1.84	45	0.82	0.04	0.47	0.63	0.53	6	13%
Writing Grades	First	1.90	42	0.73						
	Fourth	1.98	42	0.84	0.08	0.51	0.90	0.37	7	17%
Tutored, Nontested Students										
Reading Grades	First	**1.83**	30	0.75						
	Fourth	**2.13**	30	0.82	**0.30**	0.74	**2.19**	**0.04**	9	30%
Writing Grades	First	**1.85**	27	0.82						
	Fourth	**2.11**	27	0.93	**0.26**	0.59	**2.27**	**0.03**	9	33%
Nontutored Students										
Reading Grades	First	**2.56**	81	0.82						
	Fourth	**2.72**	81	0.90	**0.16**	0.68	**2.13**	**0.04**	24	30%
Writing Grades	First	2.58	73	0.83						
	Fourth	2.70	73	0.92	0.12	0.69	1.54	0.13	22	30%

(Statistically significant differences are in boldface.)

Integer numerical grades ranging from 1 to 4 are used in DCPS elementary schools:

1 = Child is not meeting the standard for this grade level; 2 = Child is approaching the standard; 3 = Child is meeting the standard; 4 = Child exceeds the standard.

F. School Attendance

Macro examined the attendance records of all second graders in the four study schools. The purpose of this analysis was to determine if tutored students had lower rates of absenteeism and tardiness than nontutored students. There was no significant difference in attendance among the groups. Table D-9 presents these results.

Table D-9: Comparison of Absenteeism and Tardiness of Tutored and Nontutored Students

	Group	N	Mean	Std	t	p
Comparison of Tutored Students and Not Tutored Students						
Days Absent	Tutored students	77	12.18	11.80	1.02	0.31
	Nontutored students	87	10.53	8.97		
Days Tardy	Tutored students	63	3.70	7.38	0.40	0.69
	Nontutored students	64	3.23	5.55		

G. Multivariate Analysis of Student Reading Improvements

1. Purpose and Overall Approach

The primary objective of DC Reads is to improve reading skills. This study provided us with an opportunity to examine students' improvements as they relate to tutoring. The examination involves determining whether:

1) Tutored students improve due to their participation in the DC Reads tutoring program
2) Specific tutoring activities are related to students' progress.

Addressing these questions will assist in future planning for the DC Reads tutoring program.

The research design for DC Reads is based on examining change in students' performance at three points in time. The first wave constitutes a baseline measure by establishing a pre-tutoring performance level for students. The second and third waves provide information on how the same students change on the various performance measures. These changes can be incorporated into an individual level growth model. Differences in these individual level growth models from one student to the next may result from student tutoring experiences or from other experiences.

Specifying adequate and explicit models for measuring growth has been offered through a few recently developed statistical techniques, e.g., Structural Equation Modeling and Hierarchical Linear Modeling (HLM). The HLM technique (Bryk A.S. & Raudenbush S. W., 1992[1]) was selected to analyze the data for which there are three waves (i.e., selected components of the Reading Performance Battery).

The advantage of HLM for examining this data set is that it can derive individual level growth curves on the outcome measures and then examine them in terms of the kinds of tutoring received. As do many regression-based techniques, HLM can control for background variables that can affect and distort the outcome. As such, it can provide an alternative to an experimental design. For instance, improvement can be compared for students of different genders, and can be estimated controlling for gender.

The model can be represented as a 2-level model in which each tutored student's learning curve is estimated as an individual growth trajectory over three points in time. The learning curve is represented by two parameters:

1) A parameter representing the initial testing (Wave 1) outcome of the student
2) A growth rate, which is the change in the outcome between subsequent testings.

In general, we posit a linear growth curve that assumes that the changes in scores are uniform between the first and second waves and the second and third waves. In addition, we also explore a more complicated growth function that contains a quadratic term. This model explicitly suggests that change in scores during the second and third period tests are larger than in those in the earlier period. In other words, most of the benefits result in the latter period.

In either case, the individual growth parameters become the outcome measures in a Level-2 model, where they may be related to some person-level demographic variables (e.g., gender) or to specific features of the tutoring program (e.g., the length of the tutoring program, writing, reading, review activities). This level 2 analysis explicitly examines variations in the growth curves across different students, and as such can provide information on what factors are associated with growth in reading skills.

2. Variables Used in the HLM for DC Reads

Table D-10 lists and defines the variables that were used in the HLM analyses for the evaluation of the DC Reads tutoring program.

Table D-10: Description of Variables Used in HLM Analyses

Three-Wave Level 1 Outcome Measures

- Percentage of Capital Letters
- Percentage of Lower Cases Letters
- Percentage of Yopp-Singer items
- Woodcock Reading Mastery Tests: Standard Scoring
  - Word Identification
  - Word Attack
  - Passage Comprehension

Three-Wave Level 1 Predictors

DUR	Length (in days) from the Wave 1 testing day to the Wave 2 testing day, or from the Wave 2 testing day to the Wave 3 testing day.

Level Two Predictors

Person-Level Predictors

BEGEND	Total length or tutoring days from starting to ending day
ATTENT_M	Average attentiveness of the tutored students (1=poor, 2=average , and 3=excellent)

Tutoring Time

TTIMETOT	Imputed total tutoring minutes per student due to missing values

Tutoring Activities

SESSION	Total number of tutoring sessions received by tutored students
REVIEWP	Percentage of review activities in the tutoring sessions
WORDP	Percentage of working on words in the tutoring sessions
CREADP	Percentage of reading activities in the tutoring sessions
WRITINGP	Percentage of writing activities in the tutoring sessions

Demographic variables

DGENDER	Dummy coding for gender (1=male; 0=female)

BEST COPY AVAILABLE

3. Mean Scores of Study Students in Three Waves of Testing

Table D-11 presents the mean scores and standard deviations for the 51 study students used in the HLM analysis.

Table D-11: Reading Performance Battery–Means for Wave 1, Wave 2 and Wave 3

Variable Name	Reading Performance Battery Test	Wave 1			Wave 2			Wave 3		
		Mean	N	Std	Mean	N	Std	Mean	N	Std
PERCAP	Percentage of Capital Letters	98.72	51	2.95	98.49	51	5.71	99.77	51	0.91
PERLOW	Percentage of Lower Case Letters	95.93	51	4.03	96.38	51	8.08	96.23	51	5.57
PERYOPPS	Percentage of Yopp-Singer items	52.67	51	28.40	72.01	51	19.63	76.56	51	15.00
WDID	Word Identification Standard Score	88.82	51	10.76	89.82	51	9.80	88.41	51	7.93
WDATTA	Word Attack Standard Score	89.75	51	12.97	91.29	51	13.87	92.29	51	11.75
PSGCMP	Passage Comprehension Standard Score	86.65	51	12.39	86.49	51	10.86	88.33	51	9.80

These data indicate the following:

Second-grade tutored students knew the alphabet letters well at the first testing (Wave 1), with an average of 98.72 percent of capital letters identified correctly and 95.93 percent of lower case letters identified correctly. A slight increase was evident after tutoring on both measures.

Students' phonemic awareness, as measured by the Yopp-Singer Test of Phonemic Segmentation, increased from 52.67 percent to 72.01 percent between Wave 1 and Wave 2 testing, and to 76.56 percent at the Wave 3 testing. The standard deviation or spread of students' scores on this measure decreased from 28.40 percent to 19.63 percent to 15 percent from Wave 1 to Wave 3. Thus, it appears that tutored students increased their phonemic awareness as a group and differences between students were reduced.

On average, students' Word Identification scores increased from Wave 1 to Wave 2, but decreased slightly from Wave 2 to Wave 3.

Word Attack scores increased from Wave 1 to Wave 3 (89.75 in Wave 1, 91.29 in Wave 2, and 92.29 in Wave 3).

Passage Comprehension scores decreased slightly from Wave 1 to Wave 2, but increased in Wave 3 (86.65 in Wave 1, 86.49 in Wave 2, and 88.33 in Wave 3).

Overall, the means for Woodcock Reading Mastery subtests were below the national average (mean=100, standard deviation=15). In examining the frequencies of scores, we found that the majority of students had scores above 85 in all waves of testing, but the percentage of students with scores above 85 increased from Wave 1 to Wave 3.

- For Word identification, 69 percent of students had scores above 85 in Wave 1, as compared with 76 percent in Wave 3.
- For Word Attack, 74 percent in Wave 1 versus 80 percent in Wave 3 attained scores greater than 85.
- For Passage Comprehension, 69 percent in Wave 1 versus 71 percent in Wave 3 attained scores greater than 85.

Across the three Woodcock Reading Mastery subtests, on average, 28 percent of the tutored students were below the national average by one standard deviation. This means that many tutored tested students were significantly behind other second graders.

Table D-12 presents the three waves of data in grade equivalent scores

Table D-12: Woodcock Reading Performance Battery Tests Means for Wave 1, Wave 2 and Wave 3–Grade Equivalent Scores

Reading Performance Battery Test	Wave 1			Wave 2			Wave 3		
	Mean	N	Std	Mean	N	Std	Mean	N	Std
Word Identification	1.74	51	0.43	2.03	51	0.46	2.20	51	0.47
Word Attack	1.68	51	1.00	2.06	51	1.18	2.36	51	1.04
Passage Comprehension	1.42	51	0.63	1.60	51	0.59	1.95	51	0.64

Before the tutoring program, these second grade students were, on average, at a 1.74 grade level in Word Identification. This means that, on average, students were able to identify words at the expected level for a first grader in the seventh month of the school year. At Wave 2 testing (about three months later), the average Word Identification score was up to 2.03 and, at Wave 3 up to 2.20. A similar pattern is seen for Word Attack scores (which went from an average of 1.68 to 2.06, and then to 2.36) and for Passage Comprehension scores (which went from 1.42 to 1.60, and then to 1.95).

Over the three Woodcock subtests, second grade tutored students changed from a 1.60 grade level before the tutoring program to a 2.17 grade level. This means that students, on average, gained approximately six months in grade equivalent during their six months of participation in the DC Reads tutoring program (and their participation in their school's reading program).

4. Descriptive Statistics for Independent Variables

Table D-13 presents the descriptive statistics for the independent variables used in the HLM analysis.

Table D-13: Description of the Independent Variables

Independent Variables	Description	Mean	N	Std
DUR1[1,2]	Days from Wave 1 to tutoring start day	10.45	51	22.06
DUR2[1]	Days from tutoring start day to Wave 2	88.16	51	21.39
DUR3	Days from Wave 2 to Wave 3	107.92	51	6.36
BEGEND	Total days from first day to last day of tutoring	173.51	51	37.61
FTIMETOT	Imputed total tutoring time per student (in minutes)	894.06	51	345.26
SESSION	Total number of tutoring sessions per student	20.00	51	7.17
REVIEWP	Percentage of sessions with review activities	78.81	51	20.53
WORDP	Percentage of sessions with working on words activities	86.56	51	22.37
CREADP	Percentage of sessions with reading activities	91.63	51	14.65
WRITINGP	Percentage of sessions with writing activities	87.30	51	14.21
ATTENT_M	Mean for attentiveness	2.51	51	0.34
DGENDER	Percentage of boys	0.57	51	0.50

[1] 2 students began tutoring after the first two test waves. Their DUR1 and DUR2 have been set to zero.

[2] 14 students began tutoring after the first test wave. Their DUR1 has been set to zero. (9 of the 14 students were tested and started the tutoring within 2-7 days. 5 of the 14 students were tested and started the tutoring within 9-14 days.)

The major demographic measure examined in this analysis is gender. In all, there were 29 males out of 51 tutored students. On average, there were approximately 10 days from the Wave 1 testing day to the beginning of tutoring and approximately 88 days from the beginning of tutoring to the Wave 2 testing day. There were, on average, approximately 108 days from the Wave 2 testing day to the Wave 3 testing day. Overall, the period from Wave 1 to Wave 3 testing is somewhat less than 7 months, with an average of 205 days.

Virtually all (50 out of 51) tutored students ended the tutoring program, and then took the Wave 3 test. There was an average of 174 days from the beginning to the end of receiving tutoring. The averaged tutoring time per session per student was 45 minutes, thus indicating that students received, on average, 736 minutes of tutoring. Due to some missing values for tutoring time on student progress logs, we imputed the total tutoring time per student. Overall, the total tutoring time per students is 894 or about 15 hours during the program.

On average, each tutored, tested student received about 20 tutoring sessions during the 1998–1999 school year. As described elsewhere in this report (most elaborately in Appendix B: Program Description), tutoring sessions included a number of activities. Approximately 92 percent of the sessions included reading aloud from a book; 87 percent included writing; 86 percent included working with words; and 79 percent included review activities. Students showed above average attention during tutoring sessions with an average attentiveness score of 2.51 on a 3-point scale (1=poor, 2=average, 3=excellent).

BEST COPY AVAILABLE

5. The Modeling Process

The first step in the modeling process was to create a base model for each outcome variable. For each outcome variable, we attempted to fit a linear and quadratic curve. Once the base model was selected, the relationship between the growth rate and tutoring activities and other program features could be investigated.

Due to the nature of the tutoring activities, it is highly possible that there is interaction or confounding effects among the predictors; thus the use of some predictors may suggest that other predictors should be dropped from the model. The first predictors examined were duration of participation in the program, imputed total hours of tutoring time, and number of sessions.

Number of sessions was not used in the model because it was used to calculate the percentage of sessions with review, working on words, reading, and writing activities. Number of sessions has a moderate correlation with duration of the program($r=.6697$, $p<=.000$) and a high correlation with total time of tutoring($r=.9252$,$p<=.000$). Thus, the number of tutoring sessions was viewed as being represented by these four other measures.

The second group of predictors was then considered for entry into the model. Predictors considered include presence of review, working with words, reading and writing activities; attentiveness; and gender. If any predictor was not significant as a single predictor, it was dropped from the model. Different combinations of those significant predictors were tested next. For this analysis, the number of iterations was set as 5,000 and the convergence criterion was set as .000001.

In the next sections, we present tables and explanation for the learning growth models on the outcome measures used in this study. For all measures, either percentage or standard scores were used.

6. Learning Growth Models

In the following sections we present linear growth models for all six outcome measures and quadratic models for some measures.

a. Percentage of Capital Letters

Table D-14: Linear Model of Growth in Percentage of Capital Letters

Fixed Effect	Coefficient	Standard Error	T-ratio	P-value
Wave I Score	98.440750	0.538257	182.888	0.000
Linear Growth Rate	0.008224	0.006575	1.251	0.217

Random Effect	Standard Deviation	Variance Component	df	Chi-square	P-value
Variation in Wave I Scores	0.32174	0.10352	50	28.24039	>.500
Variation in Linear Growth Rate	0.00127	0.00000	50	20.89582	>.500
Error	3.74618	14.03384			

The estimated initial percentage of capital letters recognized is 98.44 percent, which is a very high initial starting point. Growth in recognition was not significant during the period and there is little evidence that students varied in the percentage of Capital Letters identified in Wave I testing or in their growth across the three waves of testing. This result probably reflects the initial high level of recognition among students and the lack of room for improvement.

b. Percentage of Lower Case Letters

Table D-15: Linear Model of Growth in Percentage of Lower Case Letters

Fixed Effect	Coefficient	Standard Error	T-ratio	P-value
Wave I Score	95.820347	0.819837	116.877	0.000
Linear Growth Rate	0.005077	0.009183	0.553	0.582

Random Effect	Standard Deviation	Variance Component	df	Chi-square	P-value
Variation in Wave I Scores	2.59094	6.71299	50	35.75492	>.500
Variation in Linear Growth Rate	0.01161	0.00013	50	29.38436	>.500
Error	5.11634	26.17694			

The estimated initial percentage of lower case letters identified correctly is 95.82 percent across all students and growth was not significant from Wave 1 to Wave 3. The indication is that students did not significantly vary in the percentage of Lower Case Letters identified in Wave 1 testing. There is also no evidence to indicate that there was significant variation in their growth rates. Again these results probably reflect the inability of individuals to grow in these skills due to the overall high level of skill already present.

c. Percentage of Yopp-Singer Items

Table D-16: Linear Model of Growth in Phonemic Awareness

Fixed Effect	Coefficient	Standard Error	T-ratio	P-value	
Wave 1 Score	51.299162	4.113301	12.472	0.000	
Linear Growth Rate	0.233507	0.038894	6.004	0.000	
Random Effect	**Standard Deviation**	**Variance Component**	**df**	**Chi-square**	**P-value**
Variation in Wave 1 Scores	25.78028	664.62276	50	226.34721	0.000
Variation in Growth	0.21733	0.04723	50	129.64140	0.000
Error	13.54404	183.44109			

Correlation between Wave 1 and growth rate = -0.935
Reliability of Coefficient Estimate: Wave 1= .768, growth rate= .609

The average phonemic awareness at Wave 1 was 51.30 percent and the students gained an average of .23 percent per day from Wave 1 to Wave 3 testing. There is significant variation in students' Wave 1 phonemic awareness and in their learning rate.

The correlation between the Wave 1 status and growth rate is -0.935. This means students who had limited phonemic awareness at entry into the tutoring program tended to gain at a much faster rate.

The estimated reliabilities for Wave 1 scores and learning rates were .768 and .609.

Tutoring variables and gender were also considered for entry into the model. However, these predictors were not significant at the first step and so were not used in the final estimation.

d. Woodcock Word Identification

Table D-17: Quadratic Model of Word identification

Fixed Effect	Coefficient	Standard Error	T-ratio	P-value
Wave I Score	88.515578	1.545672	57.267	0.000
Linear Growth Rate	0.067981	0.032477	2.093	0.041
Quadratic Growth Rate	-0.000608	0.000288	-2.115	0.039

Random Effect	Standard Deviation	Variance Component	df	Chi-square	P-value
Variation in Wave I Scores	10.55866	111.48523	50	881.69560	0.000
Variation in Linear Growth Rate	0.03249	0.00106	50	95.22428	0.000
Error	2.67518	7.15660			

Correlation between Wave 1 and growth rate = -0.776
Reliability of Coefficient Estimate: Wave 1= 934, growth rate=.473

The estimated initial Word Identification score among students is 88.52 (national average=100). There was significant variation among students in Word Identification scores at Wave 1. There was also significant growth among students during the period; however, the growth rate itself seemed to be decreasing (as evidenced by the second quadratic growth term) as students were exposed to more tutoring. This may mean that improvement on this measure was largely confined to the first tutoring period. There was also significant variation in the linear growth parameter. Students who had limited Word Identification at entry into tutoring program tended to gain at a faster growth rate (r=-0.776).

The average growth rate dropped to .044 after 3 months and to 0.16 after 6 months. The standard deviation of the individual observations around individual growth curves was 2.68.

Table 18: Quadratic Model of Growth in Word Identification (Effects of Reading Activities)

Fixed Effect	Coefficient	Standard Error	T-ratio	P-value
Wave I Score	88.572050	1.552567	57.049	0.000
Linear Growth Rate	0.058920	0.032777	1.798	0.078
Quadratic Growth Rate	-0.0000526	0.0000290	-1.815	0.075
Effect of % Sessions With Reading on Quadratic Growth Rate	0.0000006	0.0000003	2.082	0.042

Random Effect	Standard Deviation	Variance Component	df	Chi-square	P-value
Variation in Wave I Scores	10.61227	112.62028	50	896.36453	0.000
Variation in Linear Growth Rate	0.03520	0.00124	50	101.16404	0.000
Error	2.66017	7.07649			

When the percentage of sessions that included reading activities was introduced into the model, we found that it was positively related to the quadratic growth term in Word Identification. Overall, the higher the percentage of sessions that included reading activities (i.e., reading aloud from a lesson or choice book), the higher the growth rate.

e. Woodcock Word Attack

Table D-19: Linear Model of Growth in Word Attack

Fixed Effect	Coefficient	Standard Error	T-ratio	P-value
Wave I Score	89.450362	1.896719	47.161	0.000
Linear Growth Rate	0.024865	0.012955	1.919	0.060

Random Effect	Standard Deviation	Variance Component	df	Chi-square	P-value
Variation in Wave I Scores	11.30112	127.71533	50	166.19494	0.000
Variation in Linear Growth Rate	0.01165	0.00014	50	38.66920	>.500
Error	7.24512	52.49179			

The estimated score on Word Attack at Wave 1 is 89.45. We did not find evidence of significant growth from Wave 1 to Wave 3 (p=.06). While there was significant variation in students' Wave 1 scores on Word Attack, there is no evidence to support that there was significant variation in their learning rates.

f. Woodcock Passage Comprehension

Table D-20: Linear Model of Growth in Passage Comprehension

Fixed Effect	Coefficient	Standard Error	T-ratio	P-value	
Wave 1 Score	86.282037	1.809349	47.687	0.000	
Linear Growth Rate	0.014302	0.013355	1.071	0.290	
Random Effect	**Standard Deviation**	**Variance Component**	**df**	**Chi-square**	**P-value**
Variation in Wave 1 Scores	11.83811	140.14079	50	332.99918	0.000
Variation in Linear Growth Rate	0.07092	0.00503	50	112.10278	0.000
Error	4.97829	24.78334			

Correlation between Wave 1 and growth rate = -0.684

Reliability of Coefficient Estimate: Wave 1 status=0.838, growth rate=0.552

The average score on Passage Comprehension at Wave 1 was 86.28. There was significant variation in students' Passage Comprehension scores in Wave 1. While the overall growth rate is not significant, there are individual differences in learning rate. This would indicate that growth curves are very different across individuals with positive and negative growth rates canceling each other out.

Students who had limited Passage Comprehension at entry into the tutoring program tended to gain at a faster rate (r=-0.684). The estimated reliabilities for Wave 1 scores and learning rates are .838 and .552.

We introduced two predictors into the model to explain the individual difference in growth rate. Entered predictors are total length of the tutoring program and the percentage of sessions that included reading activities.

Table D-21: Linear Model of Growth in Passage Comprehension (Effects of Total Days in the DC Reads Program and Percentage of Reading Activities)

Fixed Effect	Coefficient	Standard Error	T-ratio	P-value
Wave 1 Scores	86.260912	1.806431	47.752	0.000
Linear Growth Rate	0.015534	0.013256	1.172	0.247
Effect of Duration of Tutoring on Linear Growth Term	-0.000616	0.000253	-2.433	0.019
Effect of % Sessions with Reading on Linear Growth Term	0.001545	0.000630	2.451	0.018

Random Effect	Standard Deviation	Variance Component	df	Chi-square	P-value
Variation in Wave 1 Scores	11.81789	139.66258	50	333.36955	0.000
Variation in Linear Growth Rate	0.06983	0.00488	48	108.28151	0.000
Error	4.97538	24.75439			

Both total number of days in the tutoring program and the percentage of sessions that included reading activities are related to the learning rate on Passage Comprehension. Total number of days is related negatively and number of sessions with reading activities is related positively.

Based on the above results, we conclude the following:

For percentage of capital and lower case letters identified correctly, there is no evidence to support any growth or any individual difference in growth rates. However, initial scores were uniformly high and left little room for growth.

There is individual difference in Word Attack at Wave 1, but no evidence of growth or any individual difference in growth rates.

There are significant results for percentage of Yopp-Singer items and for Woodcock Word Identification and Passage Comprehension. The percentage of sessions that included reading is positively related to the growth rate of Word Identification. The percentage of sessions that included reading is positively related to growth in Passage Comprehension; total number of days in the program is negatively related to it.

Students who began the program with lower scores grew at a faster rate on three measures: the Yopp-Singer Test of Phonemic Segmentation and the Woodcock Word Identification and Passage Comprehension subtests

There appear to be other factors that are affecting students' performance. Even after the various predictors have been introduced into the model, there is still significant variation left in the linear growth rates for phonemic awareness and Passage Comprehension.

H. Summary of Findings

Based on the above discussion of the analyses, we draw the following conclusions.

Because DC Reads targets low-performing students, students were selected purposively for tutoring. Our analyses of the Reading Performance Battery tests indicate that teachers did select low performers and that tutored students randomly selected for our study did not differ from other tutored students. Due to individualized testing, the tester difference was also examined for each period of data collection, and there is no evidence of tester difference that might have biased results.

Overall, tutored students received about 20 sessions (approximately 15 hours) over a six-month period. This is about one session per week on average. Sessions lasted an average of 45 minutes. The most common activity in the sessions was reading aloud from a lesson or choice book. For a child who received 20 sessions, typically 18 of 20 would include reading, 17 or 20 would include working with words, 17 of 20 would include writing, and 16 of 20 would include review activities.

Students entered DC Reads with positive attitudes about reading. There is no evidence to support the notion that tutored, tested students' attitudes toward reading became more positive. This suggests that either attitudes were very positive to begin with, or that the attitude measure did not effectively discriminate different levels of attitude.

The second grade, tutored students who entered DC Reads tutoring knew the alphabet letters very well even in Wave 1 testing. This left little room for growth. For identification of Capital Letters and Lower Case Letters, there is no evidence to support any growth or individual variation in growth rates.

Between Wave 1 and Wave 3, students showed statistically significant increases in performance on the following measures: Yopp-Singer Test of Phonemic Segmentation, Woodcock Word Attack, Durrell Oral Reading Time, and Durrell Oral Reading Comprehension. They did not improve significantly in Word Identification or Passage Comprehension. A more detailed analysis showed that children who entered DC Reads with low skills in phonemic awareness, children with low skills in word identification, and children with low skills in passage comprehension tended to gain at a faster rate. The frequency of reading activities was positively associated with growth in Word Identification and Passage Comprehension. These findings support the conclusion that DC Reads

contributed to improvements in some aspects of reading performance, especially for the lowest performing children.

Overall, for the three Woodcock Reading Mastery Test subtests used, the second grade, tutored, tested students changed from about 1.60 grade level to about 2.17 grade level between Wave 1 and Wave 3 testing. This means that they gained approximately six months in reading skills during the six months they participated in DC Reads.

---

[1] Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical Linear Models: Applications and Data Analysis Methods*. Newbury Park CA: Sage.

Acknowledgements

Macro International Inc. was pleased to have the opportunity to evaluate the DC Reads program during the 1998-1999 school year. The Corporation for National Service sponsored the evaluation. Dr. Susan Labin, the evaluation task manager for the study in the Corporations Office of Evaluation, provided technical oversight and offered many helpful suggestions to the evaluation team. Karen Hallerman, program manager in the Corporation's Office of Program Planning and Integration, was very interested in and supportive of the research effort.

DC Reads is led by a collaborative of local public and private organizations, higher education institutions, and the District of Columbia Public Schools. Throughout the evaluation, Macro received the cooperation of the DC Reads partners, particularly Monica Swann, deputy director of Communities in Schools of Washington, DC; Sue Deerin, of the DC Public Schools; and Sharon Morganthaler, project director of the Collaborative Initiative for Tutoring Youth, a project of Georgetown University.

Many DC Read participants—teachers, principals, parents, Federal Work Study tutors, AmeriCorps*VISTA members, campus coordinators, and project site directors—shared their views, experience, and time with us so that we could collect quality data in this Year 2 evaluation of the program. We offer special thanks to the AmeriCorps*VISTA members, particularly those at the four primary study sites. In addition to their other responsibilities, the AmeriCorps*VISTA members at these sites helped us to collect data from parents, teachers and principals. They also facilitated individual testing of students, provided us with copies of student progress and other logs, and helped us to obtain data from school records.

We are also grateful for the assistance provided by Dr. George Farkas, director of the Center for Education and Social Policy at the University of Texas, Dallas; and by Dr. Barbara Wasik, principal research scientist at the Center for the Social Organization of Schools at Johns Hopkins University.